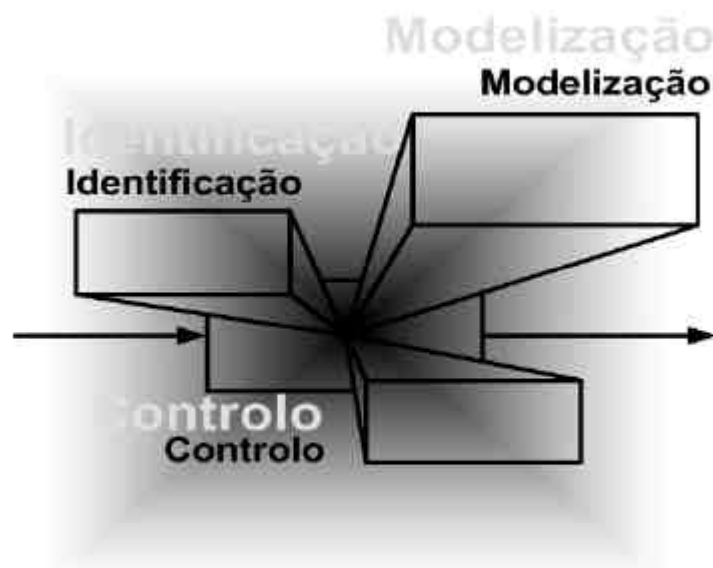


## Modelização Cinzenta Aplicada à Síntese de Controladores com Baixa Complexidade





**José António Barros  
Vieira**

**Modelização Cinzenta Aplicada à Síntese de  
Controladores com Baixa Complexidade**

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Engenharia Electrotécnica, realizada sob a orientação científica do Dr. Alexandre Manuel Mota, Professor Associado do Departamento de Engenharia Electrónica e Telecomunicações da Universidade de Aveiro

À Manuela,  
e à Sofia.

## **o júri**

presidente

Prof. Doutor Luís Machado de Abreu  
Professor Catedrático da Universidade de Aveiro

vogais

Prof. Doutor Carlos Alberto Caridade Monteiro e Couto  
Professor Catedrático da Escola de Engenharia da Universidade do Minho

Prof. Doutor Urbano José Carreira Nunes  
Professor Associado da Faculdade de Ciências e Tecnologia da Universidade de Coimbra

Prof. Doutor Alexandre Manuel Moutela Nunes da Mota  
Professor Associado da Universidade de Aveiro (orientador)

Prof. Doutor Tomás António Mendes Oliveira e Silva  
Professor Associado da Universidade de Aveiro

Prof. Doutor José António Tenreiro Machado  
Professor Coordenador com Agregação do Instituto Superior de Engenharia do Porto, Instituto Politécnico do Porto

Prof. Doutor Pedro Nicolau Faria da Fonseca  
Professor Auxiliar da Universidade de Aveiro

## **agradecimentos**

O trabalho realizado durante os últimos três anos e apresentado nesta dissertação teve a colaboração de diversas pessoas, de forma directa ou indirecta. Quero a todas elas deixar expresso o meu sincero agradecimento. No entanto, pela sua proximidade e envolvimento, queria particularizar alguns agradecimentos:

a Alexandre Mota, Professor na Universidade de Aveiro e meu orientador Científico, a quem expresso um profundo reconhecimento pelo empenho, disponibilidade, apoio e sentido crítico demonstrados na orientação deste trabalho e pela grande paciência na revisão do texto desta dissertação;

aos Engenheiros Carlos Barreto e Daniel Trigo, pela disponibilização de instalações e equipamentos da sua empresa permitindo a realização de todos os testes práticos apresentados neste trabalho e ainda pela amizade e incentivo mostrados;

a Luís Almeida, Professor na Universidade de Aveiro e meu colega de gabinete durante quase todo este período, pelas enriquecedoras e proveitosas discussões e pelo apoio e amizade demonstradas;

a Fernando Dias, da Escola Superior de Tecnologia de Setúbal, não só pelas discussões técnicas e científicas e trabalho realizado em comum, mas principalmente pela amizade evidenciada;

a Miguel Mendonça, da Escola Superior de Tecnologia e Gestão de Águeda, pelo apoio prestado na descrição do comportamento físico do esquentador;

aos Professores José Aberto Fonseca, Pedro Fonseca e Paulo Pedreiras, da Universidade de Aveiro, pelo apoio e amizade demonstradas;

a Mário Paulo Martins, pela amizade, disponibilidade e apoio na revisão do texto desta dissertação;

a todos os colegas e amigos dentro e fora do Departamento de Engenharia Electrotécnica da Escola Superior de Tecnologia de Castelo Branco e do Departamento de Electrónica e Telecomunicações da Universidade de Aveiro, pela amizade e incentivo com que sempre me apoiaram e que assim me ajudaram a chegar até aqui;

à minha Mãe, pelo apoio, incentivo e preocupação demonstradas que nunca esquecerei;

e, acima de tudo, à Manuela, pelo apoio, encorajamento e paciência tidas ao longo destes anos, e à Sofia pelo seu bom comportamento nas minhas ausências.

## palavras-chave

Modelização cinzenta, interpretabilidade *versus* complexidade de modelos, modelização neuro-difusa, modelização orientada por blocos, modelização híbrida, controlo baseado em modelo, controlo Smith preditivo e controlo de um esquentador,

## resumo

Nas últimas décadas tem-se assistido à automatização generalizada da maioria dos sistemas e equipamentos que nos rodeiam no dia-a-dia, desde os sistemas industriais, de transporte, até aos pequenos electrodomésticos. A automatização dos sistemas torna-os mais inteligentes, no sentido de maior capacidade de adaptação operacional e maior eficácia, facilitando e simplificando a sua utilização.

O problema que serviu de motivação ao desenvolvimento deste trabalho foi precisamente a automatização de um equipamento de aquecimento de água a gás, conhecido por esquentador doméstico.

Neste sentido, o presente trabalho propõe novas estratégias de automatização inteligentes para o controlo de um esquentador, adaptadas à execução em sistemas embutidos de baixo poder de cálculo.

Este trabalho suporta a tese de que a utilização de modelos com um elevado nível de interpretabilidade propicia a construção de estruturas com baixa complexidade matemática, potenciando a simplicidade das malhas de controlo. A interpretabilidade dos modelos deste tipo, tipicamente classificados como de caixa cinzenta ou cinzentos, depende do tipo de conhecimentos utilizados e incorporados na sua construção. Existindo, desde modelos construídos e identificados com base em dados de entrada e saída que apresentam baixos níveis de interpretabilidade, passando por modelos baseados em conhecimentos experimentais com níveis de interpretabilidade intermédios até modelos baseados em conhecimentos fenomenológicos que apresentam elevados níveis de interpretabilidade.

Com o objectivo de construir modelos para o esquentador com um razoável nível de interpretabilidade são neste trabalho apresentados e explorados três métodos de modelização neuro-difusa (modelização cinzenta escura) orientada por blocos (modelização cinzenta) e híbrida (modelização cinzenta clara). Neste âmbito, são propostos como modelos para o esquentador: o modelo neuro-difuso, Hammerstein adaptativo polinomial/neuro-difuso e híbrido série polinomial/neuro-difuso.

Utilizando os referidos modelos são propostas três malhas de controlo baseadas no controlador Smith preditivo, com algumas simplificações que resultam da interpretabilidade dos respectivos modelos, tal como a linearização do esquentador relativamente ao fluxo de gás.

Finalmente, os desempenhos dos vários controladores assim como os respectivos níveis de interpretabilidade e de complexidade matemática são comparados.

Este trabalho apresenta diversas contribuições ao nível das diferentes abordagens de modelização e da definição de algoritmos de controlo a aplicar num esquentador doméstico:

- Definição e identificação do modelo neuro-difuso, dos modelos Hammerstein adaptativo polinomial/neuro-difuso e dos modelos híbridos série polinomial/neuro-difuso, que exploram as várias metodologias de combinação de conhecimentos *a priori* na sua construção (modelização cinzenta);

- Proposta de definição da equação de balanço de energias de um esquentador doméstico;
- Proposta de alteração da malha de controlo Smith preditivo de forma a linearizar o sistema relativamente a uma dada variável de entrada;
- Apresentação da malha de controlo Smith preditivo para sistemas com múltiplas entradas com diferentes tempos mortos;
- Aplicação dos vários modelos cinzentos no controlo automático de um esquentador doméstico, obtendo controladores adaptativos e não adaptativos de complexidade matemática relativamente baixa.

Este trabalho disponibiliza uma gama diversificada de soluções de controlo automático para o esquentador com semelhantes níveis de desempenho, mas com diferentes níveis de interpretabilidade e de complexidade matemática. De referir que os testes e as avaliações das soluções de controlo apresentadas, além de simuladas antecipadamente num computador, foram efectuados com dados reais.

## keywords

Nonlinear modelling and control, neuro-fuzzy modeling, block oriented modeling, hybrid modeling, model based control, Smith predictive control, variable time delay and gas water heater.

## abstract

In the last decade, the number of systems and pieces of equipment that surround us in our daily life and are controlled automatically grew tremendously, from industrial systems to airplanes, cars, and even small domestic appliances. Such automatic control makes those systems *more intelligent* in the sense of higher adaptability to different operational conditions and higher efficiency, facilitating and simplifying their use.

The problem that motivated this work was one such case, namely the automatic control of a gas water heating equipment, known as domestic gas heater, for which we propose new *intelligent* control strategies particularly adapted to be embedded in low processing power systems.

Our thesis states that using models with a high level of interpretability leads to models with low mathematical complexity, thus facilitating the development of simple feedback controllers.

These models, typically called grey box models or grey models, have an interpretability that depends on the type of knowledge used and incorporated in their construction. There are models developed and identified based on input/output data that present low interpretability levels, models based on experimental knowledge with medium interpretability levels, and even models based on phenomenological knowledge that present high interpretability levels. In order to develop a model for the gas water heater with a reasonably high interpretability level we present and explore three modelling techniques: the neuro-fuzzy modelling (also called dark grey modelling); the block based modelling (grey modelling); and the hybrid modelling (also referred to as clear grey modelling). From these, we propose the following models for the gas heater: a neuro-fuzzy model, a polynomial/neuro-fuzzy adaptive Hammerstein model and a polynomial/neuro-fuzzy series hybrid model.

Using the models referred above, we propose three Smith predictive controllers with a few simplifications that arose from the models interpretability such as the gas heater linearization with respect to the gas flow. Finally, the performance of the three controllers are compared among them as well as the respective levels of interpretability and mathematical complexity.

This work presents several contributions in the area of modelling and definition of control algorithms applicable to the control of domestic gas water heaters:

- Definition and identification of the neuro-fuzzy model, polynomial/neuro-fuzzy adaptive Hammerstein models and polynomial/neuro-fuzzy series hybrid models, exploring several methodologies to combine *a priori* knowledge in their construction (grey modelling);
- Proposal for the definition of the energy balance equations in a domestic gas water heater;
- Adaptation of the Smith predictive controller on order to linearize the system with respect to a given input;
- Formulation of the Smith predictive controller for systems with multiple inputs and different time delays.

Beyond the referred contributions, we also present three controllers applied to the domestic gas water heater, one adaptive and two not adaptive, achieving very good results.



Despite their similar performance, these control solutions exhibit different levels of interpretability and mathematical complexity as expected. The controllers were previously simulated in a computer but their experimental assessment was carried out with real data.

## **apoios**

Este trabalho foi apoiado pelas seguintes instituições e empresas:

Escola Superior de Tecnologia de Castelo Branco, que me dispensou de serviço docente durante três anos e que financiou a minha participação em várias conferências internacionais;

Ministério da Educação e FSE, no âmbito do III Quadro Comunitário de Apoio, através do programa PRODEP III, que financiou parcialmente a minha dispensa de serviço docente, bem como a minha participação em várias conferências internacionais;

Departamento de Electrónica e de Telecomunicações da Universidade de Aveiro, que me apoiou financeiramente na participação em várias conferências internacionais e pelas condições proporcionadas para o desenvolvimento do trabalho realizado no âmbito desta tese;

Fundação da Ciência e Tecnologia, programa Fundo de Apoio à Comunidade Científica, pelo apoio financeiro concedido com vista à participação em algumas conferências internacionais;

TetraCis - Electrónica e Telecomunicações, Lda, que me facultou as suas instalações e equipamentos permitindo a realização de todos os testes práticos efectuados no esquentador doméstico.

# **MODELIZAÇÃO CINZENTA APLICADA À SÍNTESE DE CONTROLADORES COM BAIXA COMPLEXIDADE**

José António Barros Vieira  
Universidade de Aveiro  
TESE DE DOUTORAMENTO

15 de Janeiro de 2006



# Conteúdo

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Motivação . . . . .	1
1.2	Tese . . . . .	2
1.3	Abordagens e Enquadramento . . . . .	2
1.4	Controlo Automático de Esquentadores . . . . .	4
1.5	Estrutura desta Dissertação . . . . .	5
1.6	Contribuições desta Dissertação . . . . .	8
<b>2</b>	<b>Modelização e Identificação de Sistemas Não Lineares</b>	<b>11</b>
2.1	Introdução . . . . .	11
2.2	Modelização de Sistemas . . . . .	12
2.2.1	O que é um Modelo? . . . . .	12
2.2.2	Para que Serve um Modelo? . . . . .	12
2.2.3	Conhecimentos Utilizados na Construção de Modelos . . . . .	14
2.2.4	Tipos de Modelos . . . . .	15
2.2.5	Modelos do Tipo Caixa Cinzenta . . . . .	16
2.3	Identificação de Sistemas . . . . .	22
2.3.1	Ciclo de Identificação . . . . .	22
2.3.2	O Sistema . . . . .	24
2.3.3	Condição Experimental . . . . .	25
2.3.4	Seleção da Estrutura do Modelo . . . . .	26
2.3.5	Critérios de identificação . . . . .	29
2.3.6	Critérios de Validação . . . . .	31
2.4	Conclusão . . . . .	32
<b>3</b>	<b>Modelização e Identificação Neuro-Difusa</b>	<b>35</b>
3.1	Introdução . . . . .	35
3.2	Sistemas Difusos . . . . .	36
3.2.1	Conjuntos Difusos . . . . .	37
3.2.2	Fusificação . . . . .	39
3.2.3	Base de Conhecimentos . . . . .	39
3.2.4	Inferência Difusa . . . . .	40
3.2.5	Desfusificação . . . . .	41
3.2.6	Comentários . . . . .	42
3.3	Redes Neurais Artificiais . . . . .	43

3.3.1	Rede Neuronal Multi-Camada . . . . .	44
3.3.2	Método de Retropropagação do Erro . . . . .	45
3.3.3	Comentários . . . . .	46
3.4	Sistemas Neuro-Difusos . . . . .	47
3.4.1	Construção de um Sistema Neuro-Difuso . . . . .	47
3.4.2	Estrutura e Parâmetros de um Sistema Neuro-Difuso . . . . .	51
3.5	Identificação de Sistemas Neuro-Difusos . . . . .	56
3.5.1	Aprendizagem Baseada no Gradiente Descendente . . . . .	57
3.5.2	Aprendizagem Híbrida . . . . .	60
3.6	Sistemas Neuro-Difusos Mais Comuns . . . . .	62
3.7	Conclusão . . . . .	63
<b>4</b>	<b>Modelização e Identificação Orientada por Blocos</b>	<b>65</b>
4.1	Introdução . . . . .	65
4.2	Modelização Orientada por Blocos . . . . .	65
4.2.1	Modelização Linear . . . . .	66
4.2.2	Modelização Hammerstein . . . . .	67
4.2.3	Modelização Wiener . . . . .	69
4.3	Identificação de Modelos Orientados por Blocos . . . . .	70
4.3.1	Identificação do Bloco Linear Dinâmico Genérico . . . . .	71
4.3.2	Identificação dos Modelos Hammerstein . . . . .	72
4.3.3	Identificação dos Modelos Wiener . . . . .	75
4.4	Conclusão . . . . .	77
<b>5</b>	<b>Modelização e Identificação Híbrida</b>	<b>79</b>
5.1	Introdução . . . . .	79
5.2	Modelização Híbrida . . . . .	80
5.2.1	Definição do Modelo Fenomenológico . . . . .	81
5.2.2	Modelo Híbrido Série . . . . .	82
5.2.3	Modelo Híbrido Paralelo . . . . .	83
5.3	Identificação de Modelos Híbridos . . . . .	85
5.3.1	Identificação do Modelo Híbrido Série . . . . .	85
5.3.2	Identificação do Modelo Híbrido Paralelo . . . . .	88
5.4	Conclusão . . . . .	89
<b>6</b>	<b>Estruturas de Controlo Baseadas em Modelos</b>	<b>91</b>
6.1	Introdução . . . . .	91
6.2	Controlo com Modelo Inverso . . . . .	91
6.3	Controlo Aditivo . . . . .	93
6.4	Controlo Baseado no Modelo Interno . . . . .	95
6.4.1	Controlo Smith Preditivo Baseado no Modelo . . . . .	96
6.5	Controlo Preditivo Generalizado . . . . .	97
6.6	Conclusão . . . . .	101

<b>7</b>	<b>O Esquentador Doméstico</b>	<b>103</b>
7.1	Introdução . . . . .	103
7.2	Estrutura Física . . . . .	103
7.2.1	Elementos Mecânicos . . . . .	104
7.2.2	Sensores . . . . .	105
7.2.3	Actuadores . . . . .	106
7.3	Definição das Entradas e Saídas . . . . .	107
7.4	Gamas de Funcionamento . . . . .	107
7.5	Estados de Funcionamento . . . . .	109
7.6	Caracterização Estática e Dinâmica . . . . .	109
7.6.1	Funcionamento Estático . . . . .	111
7.6.2	Funcionamento Dinâmico . . . . .	113
7.6.3	Comentários . . . . .	118
7.7	Especificações de Desempenho . . . . .	118
7.8	Plataforma de Testes . . . . .	119
7.9	Placa de Controlo . . . . .	121
7.10	Conclusão . . . . .	124
<b>8</b>	<b>Modelização e Identificação do Esquentador</b>	<b>127</b>
8.1	Introdução . . . . .	127
8.2	Aquisição de Dados . . . . .	128
8.2.1	Escolha do Período de Amostragem . . . . .	128
8.2.2	Dados para a Identificação . . . . .	129
8.3	Modelo do Esquentador . . . . .	130
8.3.1	Modelo dos Tempos Mortos . . . . .	131
8.3.2	Modelo do Esquentador sem Tempo Morto . . . . .	134
8.3.3	Modelo Global do Esquentador . . . . .	134
8.4	Estruturas de Treino . . . . .	135
8.4.1	Identificação do Modelo Directo . . . . .	135
8.4.2	Identificação do Modelo Inverso . . . . .	136
8.5	Modelização e Identificação Linear . . . . .	138
8.5.1	Modelo Linear . . . . .	138
8.5.2	Identificação do Modelo Linear . . . . .	139
8.5.3	Resultados . . . . .	139
8.5.4	Comentários . . . . .	140
8.6	Modelização e Identificação Neuro-Difusa . . . . .	141
8.6.1	Modelo Neuro-Difuso . . . . .	142
8.6.2	Identificação do Modelo Neuro-Difuso . . . . .	144
8.6.3	Resultados . . . . .	144
8.6.4	Comentários . . . . .	145
8.7	Modelização e Identificação Orientada por Blocos . . . . .	146
8.7.1	Modelo Hammerstein . . . . .	146
8.7.2	Identificação do Modelo Hammerstein Adaptativo . . . . .	149
8.7.3	Resultados . . . . .	152
8.7.4	Comentários . . . . .	154

8.8	Modelização e Identificação Híbrida . . . . .	155
8.8.1	Modelo Físico . . . . .	156
8.8.2	Modelo Híbrido Série . . . . .	159
8.8.3	Identificação do Modelo Híbrido Série . . . . .	161
8.8.4	Resultados . . . . .	163
8.8.5	Melhoramentos ao Modelo Híbrido Série . . . . .	164
8.8.6	Comentários . . . . .	166
8.9	Comparação de resultados . . . . .	167
8.10	Conclusão . . . . .	167
<b>9</b>	<b>Controlo do Esquentador</b>	<b>169</b>
9.1	Introdução . . . . .	169
9.2	Condições de Teste . . . . .	170
9.3	Escolha do Controlador . . . . .	171
9.4	Controlador Smith Preditivo . . . . .	172
9.5	Controlador Smith Preditivo Linear . . . . .	174
9.5.1	Malha de Controlo . . . . .	174
9.5.2	Resultados . . . . .	175
9.5.3	Comentários . . . . .	177
9.6	Controlador Smith Preditivo Neuro-Difuso . . . . .	177
9.6.1	Malha de Controlo . . . . .	177
9.6.2	Resultados . . . . .	177
9.6.3	Comentários . . . . .	178
9.7	Controlador Smith Preditivo Hammerstein . . . . .	179
9.7.1	Malha de Controlo . . . . .	179
9.7.2	Resultados . . . . .	181
9.7.3	Comentários . . . . .	182
9.8	Controlador Smith Preditivo Híbrido . . . . .	184
9.8.1	Malha de Controlo . . . . .	184
9.8.2	Resultados . . . . .	185
9.8.3	Comentários . . . . .	187
9.9	Comparação de Resultados . . . . .	188
9.10	Conclusões . . . . .	189
<b>10</b>	<b>Conclusões e Trabalhos Futuros</b>	<b>191</b>
10.1	Conclusões . . . . .	191
10.2	Trabalhos Futuros . . . . .	194
<b>A</b>	<b>Lista de Publicações e Comunicações</b>	<b>209</b>
A.1	Artigos publicados em Jornal . . . . .	209
A.1.1	Nacional . . . . .	209
A.1.2	Internacional . . . . .	209
A.2	Artigos apresentados em Conferência . . . . .	210
A.2.1	Nacional . . . . .	210
A.2.2	Internacional . . . . .	210



<b>B</b>	<b>Glossário de Termos e Abreviaturas</b>	<b>213</b>
B.1	Abreviaturas . . . . .	213
B.2	Símbolos . . . . .	214
B.2.1	Genéricos . . . . .	214
B.2.2	Esquentador . . . . .	215
B.2.3	Modelos Neuro-Difusos . . . . .	216
B.2.4	Modelos Hammerstein . . . . .	216
B.2.5	Modelos Híbridos . . . . .	216
<b>C</b>	<b>Sistema de Controlo do Esquentador - <i>Hardware</i></b>	<b>217</b>
<b>D</b>	<b>Sistema de Controlo do Esquentador - <i>Software</i></b>	<b>219</b>



# Lista de Figuras

1.1	Definição da escala cinzenta dos diferentes níveis de conhecimento <i>a priori</i> utilizados na modelização e identificação de sistemas. . . . .	3
2.1	Modelo de um dado sistema. . . . .	13
2.2	Diferentes tipos de informação sobre um dado sistema que permitem a criação de diferentes tipos de modelos. . . . .	15
2.3	Diferentes aspectos e características na definição de um modelo. . . . .	15
2.4	Modelos do tipo caixa cinzenta: intercepção entre modelos do tipo caixa branca e negra. . . . .	16
2.5	Interpretabilidade fenomenológica nos vários tipos de modelos. . . . .	17
2.6	Modelo cinzento genérico (combinação dos vários tipos de conhecimento na construção do modelo final). . . . .	18
2.7	Interpretabilidade dos diferentes modelos do tipo caixa cinzenta: modelo híbrido, modelo orientados por blocos e modelo neuro-difuso. . . . .	18
2.8	Sistema difuso base da modelização neuro-difusa. . . . .	19
2.9	Modelização orientada por blocos - a) modelo Hammerstein b) modelo Wiener. . . . .	20
2.10	Modelização híbrida a) modelo paralelo b) modelo série. . . . .	22
2.11	Identificação do modelo do sistema através do sinal $e(k)$ . . . . .	23
2.12	Ciclo de identificação. . . . .	24
2.13	Estrutura paramétrica série/paralelo. . . . .	29
3.1	Estrutura de um sistema difuso. . . . .	37
3.2	Definição dos conjuntos difusos da variável temperatura. . . . .	38
3.3	Aplicação do método de desfusão do centro de área a sistemas tipo Mandani, para o exemplo do sistema de aquecimento. . . . .	42
3.4	Rede neuronal artificial com uma camada escondida ( $p$ entradas, $n$ saídas e $m$ neurónios na camada escondida). . . . .	44
3.5	Sistema neuro-difuso do tipo Mandani. . . . .	48
3.6	Sistema neuro-difuso do tipo Takagi-Sugeno. . . . .	50
3.7	Partição uniforme do espaço de entrada (para o caso de um modelo com 2 entradas e funções de pertença trapezoidais). . . . .	53
3.8	Partição não uniforme do espaço de entrada (para o caso de 2 entradas com funções de pertença trapezoidais). . . . .	54

3.9	Partição dispersa do espaço de entrada (para o caso de 2 entradas com funções de pertença trapezoidais). . . . .	55
3.10	Sistema neuro-difuso do tipo Takagi-Sugeno de ordem zero com $p$ entradas e uma saída. . . . .	57
4.1	Modelo Hammerstein. . . . .	67
4.2	Modelo Wiener. . . . .	69
5.1	Modelo fenomenológico. . . . .	81
5.2	Modelo híbrido série. . . . .	82
5.3	Modelo híbrido paralelo. . . . .	84
5.4	Identificação do modelo híbrido série baseada no erro entre $\mathbf{z}$ e $\hat{\mathbf{z}}$ . . . .	86
5.5	Identificação do modelo híbrido série baseada no erro entre $\mathbf{y}$ e $\hat{\mathbf{y}}$ . . . .	86
5.6	Identificação do modelo híbrido paralelo. . . . .	88
6.1	Controlador com modelo inverso. . . . .	91
6.2	Controlador com modelo inverso. . . . .	92
6.3	Controlador com modelo inverso para sistemas de 1 <sup>a</sup> ordem. . . . .	93
6.4	Controlador com modelo inverso para sistemas de 2 <sup>a</sup> ordem. . . . .	93
6.5	Controlador aditivo. . . . .	94
6.6	Controlador aditivo puro para sistemas de 2 <sup>a</sup> ordem. . . . .	94
6.7	Controlador aditivo misto para sistemas de 2 <sup>a</sup> ordem. . . . .	95
6.8	Controlador com modelo interno. . . . .	96
6.9	Controlador com modelo interno para modelos de 2 <sup>a</sup> ordem. . . . .	96
6.10	Controlador Smith preditivo baseado no modelo interno. . . . .	97
6.11	Controlador Smith preditivo para modelos de 2 <sup>a</sup> ordem. . . . .	98
6.12	Princípio básico do controlo preditivo. . . . .	98
6.13	Princípio básico do controlo preditivo com tempo morto. . . . .	99
6.14	Controlador preditivo generalizado. . . . .	100
7.1	Componentes constituintes de um esquentador doméstico. . . . .	104
7.2	Queimador de gás . . . . .	105
7.3	Câmara de combustão e permutador . . . . .	105
7.4	Válvulas de gás. . . . .	106
7.5	Variáveis de entrada e de saída do esquentador . . . . .	107
7.6	Modelo do esquentador com actuador incluído. . . . .	108
7.7	Variáveis de entrada e de saída do esquentador (modelo simplificado). .	111
7.8	Superfície com os pontos de funcionamento estático. . . . .	112
7.9	Aumentos da temperatura em função do fluxo de gás, para caudal fixo de 8,5 l/min. (respostas em estado estacionário). . . . .	113
7.10	Aumentos da temperatura em função ao degrau ascendente e descendente de fluxo de gás para dois caudais diferentes. . . . .	114
7.11	Aumentos da temperatura da água em função ao degrau ascendente de fluxo de gás para cinco caudais fixos diferentes (pormenor do tempo morto). . . . .	114
7.12	Tempo morto teórico e tempo morto prático em função do caudal. . . .	116

7.13	Aumento da temperatura em função de degraus ascendente e descendente de caudal da água, para um fluxo de gás fixo de 65%.	117
7.14	Aumento da temperatura em função de degraus descendentes de dois caudais da água, para um fluxo de gás fixo de 65% (pormenor do tempo morto).	117
7.15	Diagrama de blocos do esquentador e suas características funcionais.	118
7.16	Esquema global da plataforma de testes.	120
7.17	Diagrama de blocos das comunicações entre a placa de controlo e o computador pessoal.	120
7.18	Fotografia da plataforma de testes.	121
7.19	Fotografia da placa de controlo.	122
7.20	Diagrama de blocos da placa de controlo do esquentador.	123
8.1	Dados de treino. De cima para baixo, gráfico 1 - caudal de água, gráfico 2 - fluxo de gás e gráfico 3 - aumento da temperatura.	130
8.2	Dados de teste. De cima para baixo, gráfico 1 - caudal de água, gráfico 2 - fluxo de gás e gráfico 3 - aumento da temperatura.	131
8.3	Dados de treino estático e dinâmico para um caudal fixo de 8 l/min (linha continua - aumento da temperatura, linha traço ponto - fluxo de gás e linha tracejada - caudal de água).	132
8.4	Dados de treino estático e dinâmico para um fluxo de gás de 90% (linha continua - aumento da temperatura, linha traço ponto - fluxo de gás e linha tracejada - caudal de água).	133
8.5	Entradas e saídas mensuráveis do esquentador.	133
8.6	Tempos mortos relativos ao fluxo de gás e ao caudal de água ( $\tau_{m1}(t)$ e $\tau_{m2}(t)$ valores teóricos contínuos - linhas contínuas e $d_1(k)$ e $d_2(k)$ valores discretos - ° e *, respectivamente).	135
8.7	Modelo dos tempos mortos variáveis do esquentador.	135
8.8	Modelo do esquentador sem tempo morto.	136
8.9	Modelo directo global do esquentador.	136
8.10	Estrutura de treino do modelo directo global do esquentador.	137
8.11	Estrutura de treino do modelo inverso global do esquentador.	138
8.12	Resultados do modelo directo linear. De cima para baixo, gráfico 1 - aumento da temperatura estimada e real (linha tracejada e contínua, respectivamente) e gráfico 2 - sinal de erro entre os dois sinais anteriores, em ordem ao tempo.	140
8.13	Resultados do modelo inverso linear. De cima para baixo, gráfico 1 - fluxo de gás estimado e real (linha tracejada e contínua, respectivamente) e gráfico 2 - sinal de erro entre os dois sinais anteriores, em ordem ao tempo.	141
8.14	Resultados do modelo directo neuro-difuso. De cima para baixo, gráfico 1 - aumento da temperatura estimada e real (linha tracejada e contínua, respectivamente) e gráfico 2 - sinal de erro entre os dois sinais anteriores, em ordem ao tempo.	145

8.15	Resultados do modelo inverso neuro-difuso. De cima para baixo, gráfico 1 - fluxo de gás estimado e real (linha tracejada e contínua, respectivamente) e gráfico 2 - sinal de erro entre os dois sinais anteriores, em ordem ao tempo. . . . .	146
8.16	Modelo Hammerstein do esquentador. . . . .	147
8.17	Figura ilustrativa do processo de recolha dos dados em estado estacionário $\mathbf{EE}^{NEE}$ com base nos dados estáticos e dinâmicos $\mathbf{E}^{NE}$ , para um caudal fixo ( $\Delta t(k)$ - linha tracejada e $f_g(k)$ - linha contínua). . . . .	150
8.18	Identificação recursiva do bloco linear do modelo Hammerstein. . . . .	151
8.19	Resultados da identificação da função não linear estática do modelo Hammerstein adaptativo polinomial directo, função estática real e estimada (linha tracejada e contínua, respectivamente). . . . .	153
8.20	Evolução da identificação dos parâmetros do modelo linear para o modelo Hammerstein adaptativo polinomial para os dados de teste. De cima para baixo, gráfico 1 - parâmetros estimados e gráfico 2 - caudal da água. . . . .	154
8.21	Resultados do modelo Hammerstein adaptativo polinomial directo. De cima para baixo, gráfico 1 - aumento da temperatura estimada e real (linha tracejada e contínua, respectivamente) e gráfico 2 - sinal de erro entre os dois sinais anteriores, em ordem ao tempo. . . . .	155
8.22	Resultados do modelo Hammerstein adaptativo polinomial inverso. De cima para baixo, gráfico 1 - fluxo de gás estimado e real (linha tracejada e contínua, respectivamente) e gráfico 2 - sinal de erro entre os dois sinais anteriores, em ordem ao tempo. . . . .	156
8.23	Modelo híbrido série do esquentador. . . . .	159
8.24	Identificação do modelo híbrido série directo. . . . .	162
8.25	Identificação do modelo híbrido série inverso. . . . .	163
8.26	Parâmetros do modelo físico linear para o modelo híbrido para os dados de treino. De cima para baixo, gráfico 1 - parâmetros do modelo físico e gráfico 2 - caudal da água. . . . .	164
8.27	Resultados do modelo híbrido série polinomial directo. De cima para baixo, gráfico 1 - aumento da temperatura estimada e real (linha tracejada e contínua, respectivamente) e gráfico 2 - sinal de erro entre os dois sinais anteriores, em ordem ao tempo. . . . .	165
8.28	Resultados do modelo híbrido série polinomial inverso. De cima para baixo, gráfico 1 - fluxo de gás estimado e real (linha tracejada e contínua, respectivamente) e gráfico 2 - sinal de erro entre os dois sinais anteriores, em ordem ao tempo. . . . .	166
9.1	Perfis de teste, de cima para baixo, gráfico 1 - $r(k)$ sinal de referência para a temperatura da água quente, gráfico 2 - $c_a(k)$ caudal de água e gráfico 3 - $t_{af}(k)$ temperatura da água fria. . . . .	170
9.2	Controlador Smith Preditivo com modelo inverso para sistemas SISO de 1ª ordem. . . . .	172

9.3	Controlador Smith preditivo do esquentador (sistema MISO que utiliza os regressores típicos utilizados num sistema de 1º ordem). . . . .	173
9.4	Controlador Smith preditivo linear constante. . . . .	175
9.5	Resultados de controlo com o CSPL. . . . .	176
9.6	Controlador Smith preditivo neuro-difuso. . . . .	178
9.7	Resultados de controlo com o CSPND. . . . .	179
9.8	Controlador Smith preditivo Hammerstein adaptativo polinomial. . . .	180
9.9	CSPHA simplificado. . . . .	181
9.10	Resultados de controlo com o CSPHA. . . . .	182
9.11	Resultados de controlo com o CSPHA (continuação). . . . .	183
9.12	Controlador Smith preditivo híbrido série polinomial. . . . .	184
9.13	Controlador Smith preditivo híbrido série polinomial simplificado. . . .	185
9.14	Resultados de controlo com o CSPHS. . . . .	186
9.15	Resultados de controlo com o CSPHS (continuação). . . . .	187





# Lista de Tabelas

1.1	Diferentes métodos de modelização/diferentes fontes de conhecimento. .	4
3.1	Operadores lógicos difusos. . . . .	41
3.2	Método híbrido de aprendizagem. . . . .	61
7.1	Gamas de funcionamento das principais variáveis. . . . .	108
7.2	Resumo das respostas do aumento da temperatura a degraus de fluxo de gás para cinco caudais diferentes. . . . .	115
8.1	MSEs obtidos com o modelo linear. . . . .	139
8.2	MSEs obtidos com o modelo neuro-difuso. . . . .	144
8.3	MSEs obtidos com o modelo Hammerstein adaptativo polinomial. . . .	152
8.4	MSEs obtidos com o modelo Hammerstein Adaptativo neuro-difuso. . .	154
8.5	MSEs obtidos com o modelo híbrido série polinomial. . . . .	163
8.6	MSEs obtidos com o modelo híbrido série neuro-difuso. . . . .	164
8.7	MSEs obtidos pelos quatro modelos: linear, neuro-difuso, Hammerstein adaptativo polinomial e híbrido série polinomial. . . . .	167
9.1	Desempenhos, sobrelevações/subelevações, tempos de estabelecimento e tempos de execução dos quatro controladores. . . . .	188
D.1	Tempos de execução e espaço em memória necessários para os quatro algoritmos de controlo. . . . .	219



# Capítulo 1

## Introdução

### 1.1 Motivação

O presente trabalho surge na sequência de um projecto de investigação e cooperação desenvolvido entre o Departamento de Electrónica e Telecomunicações da Universidade de Aveiro e a empresa TetraCis - Electrónica e Telecomunicações, Lda, cujo objectivo principal foi o desenvolvimento de algoritmos de controlo simples e robustos para o controlo automático de esquentadores de uso doméstico.

O esquentador é um equipamento a gás que permite, em tempo real, o aquecimento de um determinado caudal de água. Hoje em dia existem diversos tipos de esquentadores com diferentes características. No entanto, a maior parte destes equipamentos a gás são completamente mecânicos e não possuem um sistema de controlo efectivo da temperatura da água. Tipicamente, os esquentadores disponíveis no mercado são sistemas em malha aberta, cuja regulação da temperatura é realizada manualmente pelo utilizador. Essa regulação é efectuada junto à torneira, adicionando mais ou menos água fria à água quente proveniente do esquentador.

Este tipo de regulação apresenta para os utilizadores vários inconvenientes, dos quais se destacam o:

- desperdício de tempo (já que a cada utilização é necessário regular manualmente a temperatura da água);
- desperdício da água (uma vez que, para regular a temperatura da água é necessário adicionar água fria aumentando o consumo);
- desperdício de energia (já que, tipicamente, a água é aquecida a uma temperatura superior à temperatura desejada);
- desconforto térmico (uma vez que, após a regulação manual, ainda podem ocorrer variações na temperatura da água originadas, essencialmente, por variações que ocorrem no caudal de água provocadas pela abertura de outras torneiras).

Os inconvenientes resultantes de uma regulação manual da temperatura, a crescente exigência dos utilizadores, a competitividade entre os fabricantes e o grande desenvolvimento tecnológico e científico na área da modelização e controlo automático de sistemas

impulsionaram, nas últimas décadas, o aparecimento de soluções de controlo aplicadas ao controlo deste tipo de esquentadores. Estas soluções apresentam algoritmos de controlo executados por sistemas computacionais permitindo uma regulação automática da temperatura desejada [HCRH94], [HH95], [AMM96], [MF96], [RBQG97]. Uma vez que estes controladores são executados por sistemas computacionais embutidos de baixo poder de cálculo, as estratégias de controlo devem passar pelo desenvolvimento de algoritmos com baixa complexidade matemática.

## 1.2 Tese

O presente trabalho propõe várias estratégias de automatização inteligentes para o controlo de um esquentador doméstico adaptadas à execução em sistemas embutidos de baixo poder de cálculo. Estas estratégias passam pela exploração de algoritmos de controlo baseados em modelos do tipo caixa cinzenta ou cinzentos. Estes modelos são caracterizados por um determinado grau de interpretabilidade associado a estruturas com um determinado nível de complexidade matemática.

Neste sentido, esta dissertação suporta a tese de que:

*A utilização de modelos com um elevado nível de interpretabilidade propicia a construção de estruturas com baixa complexidade matemática, potenciando a simplicidade das malhas de controlo.*

## 1.3 Abordagens e Enquadramento

Perante um sistema não linear, como é o caso do esquentador, optou-se pela utilização de algoritmos de controlo baseados em modelos. Para tornar os referidos algoritmos computacionalmente simples e com elevado grau de interpretabilidade, optou-se pelo estudo e exploração de métodos de modelização que se baseiam nas várias fontes de conhecimentos existentes e disponíveis sobre o sistema, desde um simples conjunto de dados de entrada e saída, até às equações de balanço do sistema, passando pelos conhecimentos empíricos e experimentais. A modelização de sistemas efectuada com base nas várias fontes de conhecimento referidas é conhecida por modelização do tipo caixa cinzenta, ou mais simplesmente, modelização cinzenta.

A modelização cinzenta é um assunto relativamente novo e uma área de investigação bastante ampla e de grande potencial. Esta forma de modelização baseia-se em conhecimentos *a priori* e em dados de entrada e saída do sistema para a construção do respectivo modelo. Os conhecimentos *a priori* correspondem a todos aqueles que se podem adquirir previamente, caso dos conhecimentos empíricos ou experimentais e conhecimentos fenomenológicos.

O tipo de conhecimento *a priori* utilizado, assim como o nível de conhecimento aplicado, varia de caso para caso. Sohlberg [Soh98] apresenta, (figura 1.1) uma escala cinzenta proposta por Karplus [Kar76], definindo a região cinzenta da modelização do tipo caixa cinzenta.

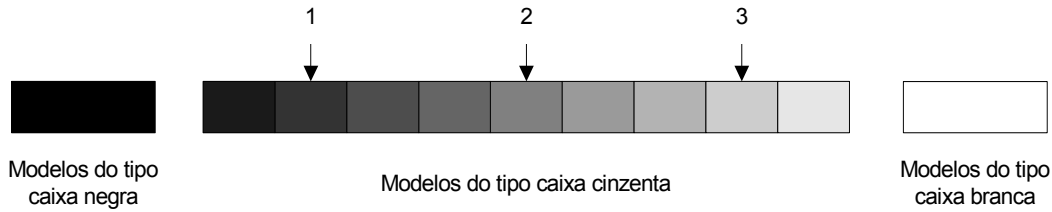


Figura 1.1: Definição da escala cinzenta dos diferentes níveis de conhecimento *a priori* utilizados na modelização e identificação de sistemas.

A tonalidade do cinzento varia de acordo com o nível e a quantidade de conhecimentos *a priori* utilizados na modelização, construindo modelos com diferentes níveis de interpretabilidade. Por exemplo, procedimentos em que os conhecimentos *a priori* se limitam à definição dos regressores do modelo, situam-se na região 1; procedimentos em que os conhecimentos *a priori* se utilizam na estimação dos parâmetros (por exemplo restrições), situam-se na região 2; procedimentos em que a estrutura é definida por leis fenomenológicas e alguns dos parâmetros são estimados com dados de entrada e saída, situam-se na região 3.

Nos últimos anos vários trabalhos citam o termo “modelização cinzenta” (*grey modelling*) e o termo “conhecimento *a priori*” na identificação de sistemas como, por exemplo [TK94b], [PU92b], [Kar76], [Soh98], [TJW96], [PP98], [JH95], [ABVS00], [SSD<sup>+</sup>94], [Cor01]. A maneira de incorporar os conhecimentos *a priori* assume várias formas, as quais irão determinar a tonalidade do cinzento do respectivo modelo.

Este trabalho explora esta abordagem de modelização, desde a construção de modelos baseados em dados de entrada e saída, modelos baseados nas equações de balanço do sistema e/ou modelos baseados em conhecimentos empíricos e experimentais. Os vários tipos de conhecimentos e as várias formas de os combinar vão conferir a estes modelos cinzentos características bastante variadas, particularmente ao nível da interpretabilidade e da complexidade matemática.

Neste sentido, este trabalho dedica especial atenção ao estudo de três técnicas de modelização e identificação cinzenta: modelização neuro-difusa (modelização cinzenta escura) [LL91], [BH91], [WM92] e [JS93]; modelização orientada por blocos (modelização cinzenta) [ZS94], [EJL91b], [NPR96] e [MAP97]; modelização híbrida (modelização cinzenta clara) [PU92b], [TK94b], [Oli98], [Lit02] e [XJ02]. O maior ou menor nível de interpretabilidade dos modelos construídos facilita a compreensão do comportamento do sistema, assim como a sua futura integração nas respectivas malhas de controlo.

A modelização neuro-difusa utiliza os conhecimentos *a priori* sobre o sistema para a definição dos regressores e das regras. Contudo, a maior fonte de conhecimentos utilizada na identificação deste tipo de modelos vem dos dados de entrada e saída. A modelização orientada por blocos utiliza os conhecimentos *a priori* sobre o sistema para a definição da divisão e da ordem dos blocos necessários, precisando ainda dos dados de entrada e saída para a identificação dos parâmetros desconhecidos. A modelização híbrida utiliza essencialmente os conhecimentos fenomenológicos sobre o sistema para

a definição da sua função de transferência, utilizando ainda os dados de entrada e saída para a identificação de alguns parâmetros desconhecidos.

A tabela 1.1 apresenta, resumidamente, as várias fontes de conhecimentos utilizadas na construção dos diferentes métodos de modelização.

Métodos de Modelização	Fontes de Conhecimento		
	Dados de Entrada/Saída	Conhecimento Empírico e Experimental	Conhecimento Fenomenoló- gico
Neuro-Difusa	X	X	-
Orientada ao Bloco	X	X	X
Híbrida	X	-	X

Tabela 1.1: Diferentes métodos de modelização/diferentes fontes de conhecimento.

O estudo e exploração destas três técnicas de modelização cinzenta levou à proposta de três modelos do esquentador com diferentes níveis de interpretabilidade e de complexidade matemática: modelo neuro-difuso, modelo Hammerstein adaptativo e modelo híbrido série. A interpretabilidade destes modelos facilita a sua integração nas malhas de controlo e permite reduzir a sua complexidade matemática.

Este trabalho apresenta três soluções de controlo baseadas nos três modelos cinzentos recorrendo à malha de controlo Smith preditiva. A complexidade matemática destas malhas de controlo depende fundamentalmente da complexidade dos respectivos modelos e controladores (modelos inversos). A utilização destes modelos no controlo automático do esquentador permite construir malhas de controlo inteligentes com elevada capacidade de adaptação, eficiência, robustez e de baixa complexidade matemática.

Esta dissertação mostra que a construção de controladores baseados em modelos cinzentos com elevado grau de interpretabilidade resulta em algoritmos de controlo de relativamente baixa complexidade matemática.

## 1.4 Controlo Automático de Esquentadores

A documentação científica de novos produtos a lançar no mercado internacional está sempre envolta em grande confidencialidade. Esta afirmação é válida para as mais recentes abordagens utilizadas no controlo automático de esquentadores domésticos. Contudo, é possível encontrar alguns estudos sobre as estratégias de controlo aplicadas em esquentadores. É evidente que muitos mais trabalhos de investigação existirão relativos ao controlo automático de esquentadores mas, por questões de confidencialidade, não estão publicados.

Dos estudos publicados sobre o controlo de esquentadores de uso doméstico, é de salientar o trabalho realizado por Claquin, apresentado em 1994 [CCR94]. Este autor define o modelo do esquentador através das suas equações diferenciais de dinâmica variável. Este modelo só foi utilizado para simulação, uma vez que necessita de variáveis teóricas de difícil ou até impossível medida no sistema. Para o controlo do esquentador, o autor propõe o controlo adaptativo baseado nos algoritmos de posicionamento de

pólos (*Pole Placement - PP*) e controlo preditivo generalizado (*Generalized Predictive Control - GPC*) utilizando um modelo de segunda ordem linear identificado em malha fechada com base no algoritmo recursivo dos mínimos quadrados (*Recursive Least Squares - RLS*).

Huyghe, em 1994 [HCRH94], e no seguimento do trabalho anterior, propõe um controlador difuso proporcional integral, apresentando um método de estimação do caudal de água. O mesmo autor, em 1995 [HH95], propõe otimizar, em malha fechada, um controlador difuso proporcional integral através do algoritmo *Simulated Annealing*.

Alves, em 1996 [AMM96], desenvolve um modelo do esquentador tendo em conta a ignição, o ruído, a quantificação da dinâmica e do tempo morto. Este modelo foi utilizado para o desenvolvimento de um controlador baseado em lógica difusa.

Silva, em 1998 [SMR98], no seguimento do trabalho de Alves, desenvolve um controlador difuso proporcional integral, otimizado através de algoritmos genéticos, no qual compensa o efeito do tempo morto existente no sistema alterando a base de regras.

Mota, em 1996 [MF96], identifica como uma razoável aproximação para o modelo do esquentador um sistema de primeira ordem variável com o caudal seguido de um tempo morto, também variável. Ao nível do controlo desenvolveu um controlador difuso proporcional integral com uma compensação aditiva proporcional à variação do caudal de água, compensando as respectivas perturbações.

Rongming, em 1997 [RBQG97], desenvolve um controlador automático misto. Baseado num controlador proporcional difuso e num controlador proporcional integral com adaptação difusa dos respectivos parâmetros. A escolha do tipo de controlador, a cada momento, depende do módulo do erro de temperatura. Com esta estratégia a compensação das cargas que actuam no sistema torna-se mais eficiente.

A maioria das soluções apresentadas possuem algoritmos relativamente complexos para serem executados por sistemas embutidos de baixo poder de cálculo e os resultados de controlo ainda apresentam erros relativamente elevados uma vez que, a maioria das soluções propostas não têm em conta a existência do tempo morto variável e a variação do caudal de água e da temperatura da água fria.

No sentido de resolver os problemas apontados esta dissertação propõem vários algoritmos de controlo com diferentes níveis de complexidade matemática e adaptabilidade apropriados para serem executados em sistemas de baixo poder de cálculo. As soluções apresentadas têm em conta as variações do tempo morto, caudal de água e temperatura da água fria do sistema.

## 1.5 Estrutura desta Dissertação

A dissertação está estruturada do seguinte modo:

### Capítulo 1 - Introdução

Neste primeiro capítulo apresentam-se a motivação, as abordagens para a resolução do problema, a estrutura da dissertação e as principais contribuições.

## Capítulo 2 - Modelização e Identificação de Sistemas Não Lineares

Apresentam-se aqui os motivos da construção de modelos, os diferentes tipos conhecimentos *a priori* utilizados na sua construção e os vários tipos de modelos. São abordados três tipos de modelização cinzenta: a neuro-difusa, a orientada por blocos e a híbrida. Apresentam-se ainda os vários passos do processo de identificação de sistemas que não se limita à escolha do tipo de modelo, efectuando outras tarefas não menos importantes, como a selecção da condição experimental, a aquisição de dados e a selecção dos critérios de identificação e validação.

## Capítulo 3 - Modelização e Identificação Neuro-Difusa

Este capítulo apresenta a modelização difusa, a modelização baseada em redes neuronais e a modelização neuro-difusa que surge da fusão das duas primeiras. Estes modelos são construídos por funções difusas ajustadas através de métodos típicos das redes neuronais construindo modelos com baixa interpretabilidade (modelos do caixa cinzenta escura). Por fim, apresentam-se os vários métodos de inicialização e identificação deste tipo de modelos, assim como, um resumo dos vários tipos de sistemas neuro-difusos desenvolvidos na última década. Do tema deste capítulo foram publicados os seguintes artigos:

- J. Vieira, F. Dias e A. Mota - Artificial Neural Networks and Neuro-Fuzzy Systems in Model and Control: A Comparative Study, *International Journal ELSEVIER, Engineering Applications of Artificial Intelligence*, 3(17):265-273, 2004.
- J. Vieira, F. M. Dias e A. Mota - Neuro-Fuzzy Systems: A Survey, *WSEAS Transactions on Systems*, 2(3): 414-419, 2004.

## Capítulo 4 - Modelização e Identificação Orientada por Blocos

Neste capítulo é apresentada a modelização orientada por blocos, mais especificamente, a modelização Hammerstein e Wiener. São ainda apresentados vários métodos de identificação dos referidos modelos, consoante o tipo de conhecimentos *a priori* existentes sobre o sistema. Estes modelos são classificados como modelos do tipo caixa cinzenta, em que a sua estrutura matemática é definida e identificada com base em conhecimentos experimentais e em dados de entrada e saída, gerando modelos com um nível de interpretabilidade razoável.

## Capítulo 5 - Modelização e Identificação Híbrida

Aborda este capítulo a modelização híbrida, apresentando a estrutura do modelo fenomenológico genérico, do modelo híbrido série e do modelo híbrido paralelo. Apresentam-se ainda, vários métodos de identificação dos referidos modelos, que dependem do modelo do tipo caixa cinzenta ou negra utilizado para a modelização da parte desconhecida do modelo fenomenológico do sistema. Esta metodologia de modelização maximiza a utilização de conhecimentos dos primeiros princípios, gerando modelos com um elevado grau de interpretabilidade, também classificados de modelos do tipo caixa cinzenta clara.



## Capítulo 6 - Estruturas de Controlo Baseadas em Modelos

Neste capítulo apresentam-se as malhas de controlo baseadas em modelos mais conhecidas.

## Capítulo 7 - O Esquentador Doméstico

Este capítulo é dedicado à apresentação pormenorizada do esquentador doméstico, da plataforma de testes desenvolvida e da electrónica da placa de controlo.

## Capítulo 8 - Modelização e Identificação do Esquentador

Neste capítulo apresenta-se os resultados da modelização e identificação do esquentador, utilizando os três métodos de modelização abordados nesta dissertação.

Relacionados com este capítulo foram publicados vários artigos. Sobre a modelização do esquentador utilizando modelos neuro-difusos publicaram-se:

- J. Vieira e A. Mota - Neural-Fuzzy Model Versus Multi-Model Based Controllers Applied in a Water Gas Heater System, *Proceedings of IASTED International Conference on Artificial Intelligence and Applications*, Benalmádena, Spain, 08 a 10 September, 1:46-51, 2003.
- J. Vieira e A. Mota - Comparison of Three Neuro-Fuzzy Architectures in Modelling a Water Gas Heater: Input Clustering Data and Input Variables Selection, *WSEAS Transactions on Systems*, 4(2):1005-1010, 2003.

Sobre a modelização orientada por blocos do esquentador foi publicado o artigo:

- J. Vieira e A. Mota - Parameter Estimation of Non-Linear Systems With Hammerstein Models Using Neuro-Fuzzy and Polynomial Approximation Approaches, *Proceedings of IEEE-FUZZ International Conference on Fuzzy Systems*, Budapest, Hungary, 25 a 29 July, 2:849-854, 2004.

Ainda relacionados com este capítulo e relativamente à modelização híbrida série do esquentador, foram publicados os seguintes artigos:

- J. Vieira e A. Mota - Combining First Principles with Grey-Box Approaches for Modelling a Water Gas Heater System, *Proceedings of IEEE 13th Mediterranean Conference on Control and Automation*, Limassol, Cyprus, 27 a 29 June, 2005.
- J. Vieira e A. Mota - Water Gas Heater Non-Linear Physical Model: Optimization with Genetic Algorithms, *Proceedings of IASTED International Conference on Modelling, Identification, and Control*, Gridelwald, Switzerland, 23 a 25 February, 1:122-127, 2004.

## Capítulo 9 - Controlo do Esquentador

Este capítulo apresenta os resultados do controlo automático do esquentador, utilizando a malha de controlo Smith preditiva e os três diferentes modelos cinzentos definidos e identificados no capítulo anterior.

Relacionados com este capítulo foram publicados vários artigos. Sobre o controlo Smith preditivo baseado em modelos neuro-difusos, foram publicados:

- J. Vieira e A. Mota - Smith Predictor Based Neural-Fuzzy Controller Applied in a Water Gas Heater that Presents a Large Time-Delay and Load Disturbances, *Proceedings IEEE International Conference on Control Applications*, Istanbul, Turkey, 23 a 25 June, 1:362-367, 2003.
- J. Vieira e A. Mota - Modelling and Control of a water Gas Heater with Neuro-Fuzzy Techniques, *Proceedings 3rd WSEAS FSFS International Conference on Fuzzy Sets and Fuzzy Systems*, Interlaken, Switzerland, 11 a 15 February, 1:56-61, 2002.

Sobre o controlo Smith preditivo do esquentador baseado no modelo Hammerstein adaptativo foi publicado o artigo:

- J. Vieira e A. Mota - Adaptive Smith Predictive Control of Non-Linear Systems Using Neuro-Fuzzy Hammerstein Models, *ICINCO-IFAC Proceedings 1st International Conference on Informatics in Control Automation and Robotics*, Setúbal, Portugal, 25 a 28 August, 3:62-69; 2004.

Relacionado ainda com este capítulo e relativamente ao controlo Smith preditivo do esquentador baseado em modelos híbridos série, foi publicado o artigo:

- J. Vieira, F. M. Dias e A. Mota - Hybrid Neuro-Fuzzy Network-Priori Knowledge Model in Temperature Control of a Gas Water Heater System, *Proceedings of the 5<sup>th</sup> International Conference on Hybrid Intelligent Systems (HIS'2005)*, Rio de Janeiro- Brasil, 06-09 November, 2005.

## Capítulo 10 - Conclusões e Trabalhos Futuros

O décimo e último capítulo apresenta as conclusões e enumera alguns dos problemas que ficaram em aberto no decorrer deste trabalho.

## 1.6 Contribuições desta Dissertação

As principais contribuições apresentadas nesta dissertação estão relacionadas com o termodoméstico estudado, mais particularmente, com o desenvolvimento de modelos do tipo caixa cinzenta do esquentador e com o controlo automático do mesmo, dentro das quais se podem destacar as seguintes:

- Exploração de várias metodologias de combinação de conhecimentos *a priori* para a construção de modelos (modelização cinzenta).
- Definição e identificação do modelo neuro-difuso, dos modelos Hammerstein adaptativo polinomial e neuro-difuso e dos modelos híbridos série polinomial e neuro-difuso para um esquentador doméstico.
- Proposta de definição da equação de balanço de energias de um esquentador doméstico.
- Proposta de simplificação da malha de controlo Smith preditivo de forma a linearizar o sistema relativamente a uma dada variável de entrada.
- Apresentação da malha de controlo Smith preditivo para sistemas com múltiplas entradas com diferentes tempos mortos.
- Aplicação dos vários modelos cinzentos no controlo automático de um esquentador doméstico, obtendo controladores adaptativos e não adaptativos de complexidade matemática relativamente baixa.



## Capítulo 2

# Modelização e Identificação de Sistemas Não Lineares

### 2.1 Introdução

A identificação de sistemas é o processo de construção de um modelo matemático de um dado sistema com base em observações e em conhecimentos adquiridos *a priori* [Nor86]. A partir desta definição, podem-se colocar várias questões, em particular, relativamente à importância e ao processo de escolha do modelo e, em geral, relativamente ao processo de identificação de sistemas.

Quanto à importância e ao processo de escolha do modelo, as questões que se podem colocar são: O que é um modelo? Para que serve? Como se constrói? Que tipo de modelos existem? No que respeita ao processo de identificação de sistemas: Que influência tem o tipo de sistema? Que género de conhecimentos sobre o sistema são necessários e podem ser utilizados? Qual o modelo que melhor representa um dado sistema? Quais os critérios de identificação e validação a utilizar? De uma forma global: Quais os passos necessários para a identificação do modelo de um sistema genérico?

A modelização e consequente identificação de sistemas pressupõem sempre uma aquisição prévia de conhecimentos, por exemplo, conhecimentos fenomenológicos, conhecimentos experimentais e conhecimentos obtidos através dos dados de entrada e saída do sistema [Mad02], [Oli98]. Com base nos referidos conhecimentos ou na conjugação destes é definida uma estrutura matemática cujos parâmetros internos são estimados de forma a minimizar um determinado critério permitindo representar o sistema real com sucesso. Neste sentido, a quantidade e qualidade de conhecimentos adquiridos, o tipo de modelo utilizado, a sua estrutura, o estabelecimento de critérios adequados de estimação de parâmetros e de validação de modelos, são factores que se revestem de grande importância no processo de identificação de sistemas [Lju87].

Na secção 2.2 define-se o que é um modelo, descreve-se a sua utilidade, qual a sua estrutura matemática e qual o tipo de conhecimentos utilizados na sua construção. Esta secção apresenta diversas formas de combinar os vários tipos de conhecimentos na construção de um dado modelo permitindo obter estruturas com diferentes graus de interpretabilidade denominadas do tipo caixa cinzenta ou cinzentas. De uma forma introdutória, são apresentados os três métodos de modelização cinzenta explorados neste

trabalho: modelização neuro-difusa, modelização orientada por blocos e modelização híbrida.

A definição da estrutura matemática do modelo é um passo muito importante no processo de identificação de sistemas, mas não o único. A secção 2.3 apresenta, de forma genérica, os vários passos do processo de identificação. Descreve os cuidados a ter em conta na selecção da condição experimental, na aquisição de conhecimentos, na selecção da estrutura matemática e nos critérios de identificação e validação do modelo.

## 2.2 Modelização de Sistemas

### 2.2.1 O que é um Modelo?

Um modelo é uma representação aproximada dos conhecimentos e comportamentos essenciais de um dado sistema.

A modelização de sistemas tornou-se uma área muito importante em qualquer domínio do conhecimento: na economia, na política, na sociologia, na saúde, na química e, de uma forma mais generalizada, no domínio da engenharia. Neste âmbito, uma definição genérica de modelo é dada por Eykhoff [Eyk74]. O autor define modelo de um sistema como sendo “uma representação dos aspectos essenciais de um dado sistema existente, ou a construir, que apresenta o conhecimento do referido sistema de forma útil”. Isto significa que o modelo é uma representação simplificada do sistema que serve de ferramenta para qualquer outra tarefa.

Dado que a modelização de sistemas é uma abstracção da realidade, o resultado desta, a maioria das vezes, é uma equação matemática. O modelo de um sistema genérico é caracterizado e definido através das respectivas variáveis de entrada e saída conhecidas e através das eventuais entradas desconhecidas, conforme ilustrado na figura 2.1. As saídas do modelo  $\hat{y}^1$  são funções das entradas  $u$  conhecidas e das entradas desconhecidas  $\xi$  (também conhecidas como perturbações), como é o caso de cargas aplicadas ao sistema, ruído nas medidas, etc.

O modelo matemático a construir deve representar o sistema o mais correctamente possível. Ao aplicar as variáveis de entrada ao modelo, este deverá fornecer variáveis de saída semelhantes às variáveis de saída do sistema.

### 2.2.2 Para que Serve um Modelo?

A construção de um modelo é raramente um objectivo por si só. O modelo é uma ferramenta que ajuda a resolver problemas que beneficiem com a descrição matemática do sistema. Aplicações de modelos no domínio das engenharias podem ser encontrados em várias áreas genéricas ([Eyk74], [Nor86], [Luy90] e [Oli98]):

---

<sup>1</sup>A utilização de  $\hat{y}$  representa a estimação da saída  $y$  de um dado sistema. Este procedimento será adoptado ao longo deste documento para distinguir a saída do modelo e a saída do sistema real designadas por uma mesma letra.

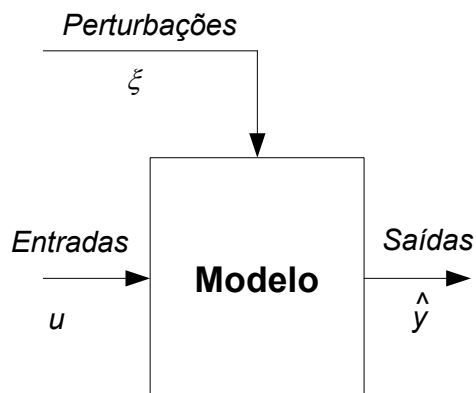


Figura 2.1: Modelo de um dado sistema.

- predição;
- supervisão;
- controlo;
- estudo do sistema (através do modelo);
- fonte de conhecimentos sobre o sistema;

Os modelos podem ser utilizados para o desenvolvimento de sensores por *software*, isto é, sensores que estimam em linha variáveis importantes as quais, por questões de custo (económico, tecnológico, ...) não podem ser medidas directamente. O conhecimento destas variáveis pode ser importante para suportar outras aplicações em linha como monitorização, controlo, predição, detecção de falhas, supervisão, etc.

A supervisão também é uma área de aplicação muito importante para os modelos, uma vez que o uso dos modelos permite a detecção e validação de falhas ocorridas nos respectivos sistemas.

Os modelos podem ser usados para o controlo de sistemas sendo neste caso designado por controlo baseado em modelos. A predição da saída dos sistemas obtida com base nos respectivos modelos é fundamental para o bom desempenho destas malhas de controlo.

Podem, igualmente, ser usados no estudo da segurança do sistema e no desenvolvimento dos controladores. A realização de experiências e testes baseados em modelos reduz drasticamente o custo económico e o tempo dispendido no desenvolvimento dos respectivos controladores.

Os modelos podem ser usados ainda para simular sistemas como, por exemplo, carros e aviões. Os simuladores permitem a realização de testes e experiências virtuais no sistema com baixo custo económico. Um caso muito conhecido é o treino de condução de carros e aviões utilizando simuladores.

### 2.2.3 Conhecimentos Utilizados na Construção de Modelos

Nos últimos anos, existe a tendência de fundir informação de natureza diversificada, e combiná-la de modo a empregar a maior quantidade de conhecimentos possível na implementação dos respectivos modelos.

A informação utilizada para a modelização e identificação de sistemas pode advir de três fontes distintas de conhecimento ([Mad02], [Oli98]):

- **Conhecimento fenomenológico do sistema.** Este conhecimento é obtido através dos chamados primeiros princípios (leis ou equações físicas, químicas, termodinâmicas, biológicas, etc.). Os modelos construídos com este tipo de conhecimentos apresentam um alto nível de sofisticação e interpretabilidade, uma vez que envolvem a compreensão e conhecimento total da cinemática dos sistemas. Os modelos construídos, unicamente, à custa deste tipo de conhecimentos são designados modelos do tipo caixa branca [Bos97];
- **Conhecimento experimental e senso comum.** Este conhecimento resulta da experiência dos operadores especializados e no senso comum aquando da operação e controlo de sistema reais. Este conhecimento é qualitativo e existe em grande quantidade, pelo menos, em ambientes industriais já instalados;
- **Conhecimento obtido através dos dados entrada e saída.** Em ambientes industriais, é tipicamente efectuado o registo das entradas e saídas dos sistemas. Estes dados contêm informação útil que pode ser utilizada na construção e identificação de sistemas. Os modelos construídos com base neste tipo de conhecimento possuem um baixíssimo grau de interpretabilidade sendo designados modelos do tipo caixa negra [Bos97].

Um sistema pode ser descrito e caracterizado através de diversos tipos de conhecimentos (como referido atrás) podendo os diferentes tipos de conhecimentos dar origem a diferentes tipos de modelos (figura 2.2).

Se existirem profundos conhecimentos fenomenológicos do sistema, a construção do modelo deve ser baseada nas respectivas equações analíticas ou fenomenológicas. Se a fonte de conhecimento do sistema é a experiência de operadores especializados que descrevem o sistema através de regras linguísticas, a construção do modelo pode ser efectuada utilizando abordagens baseadas em bases de regras. Podem ainda surgir sistemas onde a informação mais valiosa resulta dos dados de entrada e saída recolhidos durante a operação. Neste último caso, a construção do modelo é efectuada através de funções matemáticas não interpretáveis, fenomenologicamente, mas que permitam a representação do sistema.

Os modelos construídos com base na combinação destes três tipos de conhecimentos (ou da combinação de pelo menos dois) designam-se modelos do tipo caixa cinzenta. Estes podem possuir vários níveis de interpretabilidade: reduzida, caso a fonte de conhecimentos mais utilizada na construção do modelo seja o conhecimento obtido através dos dados entrada e saída, ou elevada, caso a fonte de conhecimentos mais utilizada na construção do modelo seja o conhecimento fenomenológico do sistema.



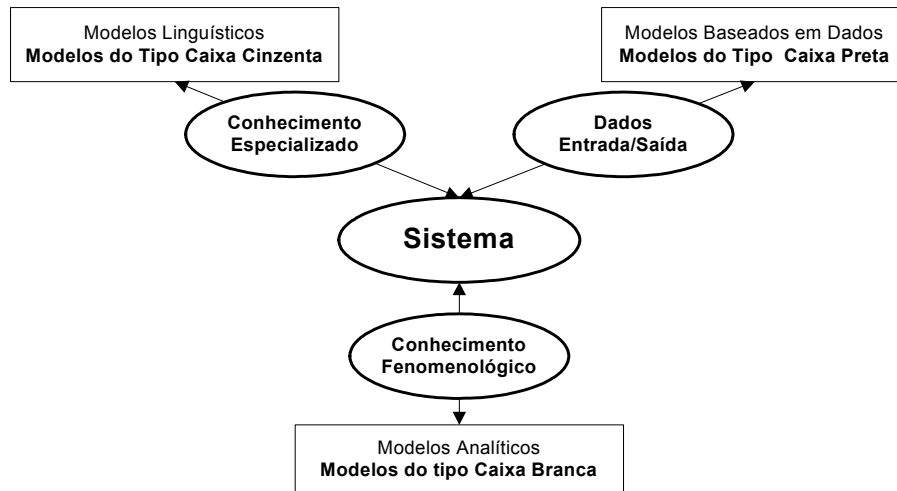


Figura 2.2: Diferentes tipos de informação sobre um dado sistema que permitem a criação de diferentes tipos de modelos.

### 2.2.4 Tipos de Modelos

Existem vários tipos de modelos, dependendo do ponto de vista em que estes são caracterizados [Lit02]. Assim, como se ilustra a figura 2.3, os modelos podem ser do tipo caixa branca ou caixa negra, lineares ou não lineares, estáticos ou dinâmicos, contínuos/analíticos ou discretos, aglomerados ou distribuídos e definidos no tempo ou na frequência.

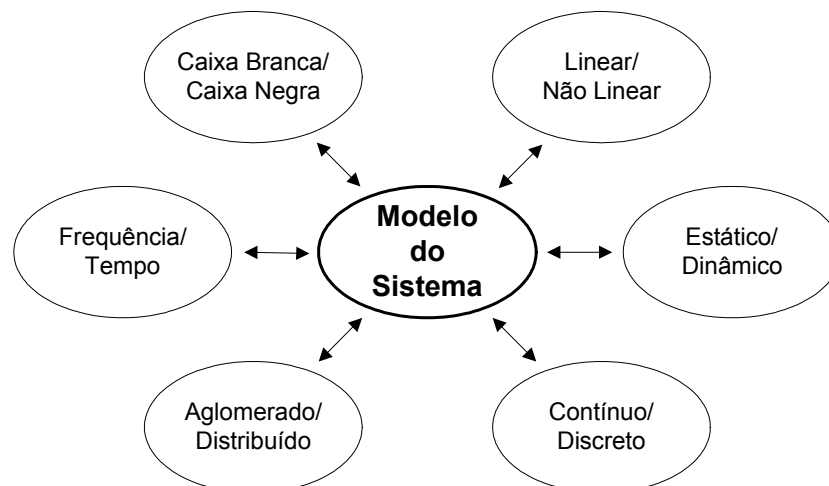


Figura 2.3: Diferentes aspectos e características na definição de um modelo.

Os modelos desenvolvidos e analisados neste trabalho são essencialmente modelos para sistemas não lineares, dinâmicos, definidos no domínio discreto e de forma concentrada explorando particularmente a interpretabilidade dos mesmos.

Os diferentes níveis interpretabilidade dos modelos levam à classificação destes em modelos do tipo caixa branca ou negra. De forma a explorar este conceito, este trabalho vai explorar os diferentes tipos de modelos, desde modelos do tipo caixa negra até modelos do tipo caixa branca, dando particular atenção aos modelos do tipo caixa cinzenta.

Os modelos do tipo caixa cinzenta são, como se ilustra na figura 2.4, um misto entre modelos do tipo caixa branca e caixa negra. Esta designação resulta essencialmente do nível de interpretabilidade que os modelos possuem e que se situa entre um nível elevado e reduzido.

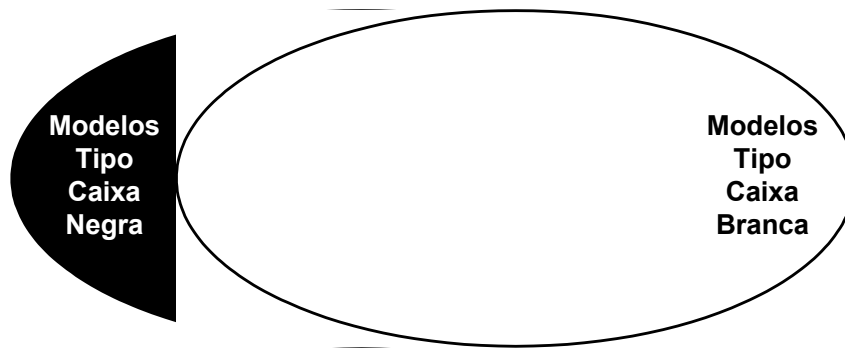


Figura 2.4: Modelos do tipo caixa cinzenta: intercepção entre modelos do tipo caixa branca e negra.

Os modelos do tipo caixa branca baseiam-se inteiramente nos conhecimentos fenomenológicos dos sistemas (equações dos balanços de massa, energia, momento, etc.). O esforço necessário para a implementação deste tipo de modelos é grande, principalmente para sistemas complexos, uma vez que estes conhecimentos podem existir de forma insuficiente ou incompleta. Em contrapartida, estes modelos descrevem o sistema com detalhe e representam-no com elevada interpretabilidade em qualquer ponto de funcionamento.

Os modelos do tipo caixa negra não possuem uma estrutura matemática que reflecta conhecimentos fenomenológicos do sistema. Possuem sim uma estrutura matemática que reflecte simplesmente a relação entre as entradas e saídas do sistema, sendo tipicamente utilizados no caso em que os únicos conhecimentos disponíveis são dados de entrada e saída. Estas estruturas matemáticas são implementadas recorrendo aos modelos auto-regressivos (como, por exemplo, modelos *Auto Regressive with eXogenous inputs* - *ARX* e *Nonlinear Auto Regressive with eXogenous inputs* - *NARX*), às *splines*, aos polinómios, às séries de *Fourier*, às redes neuronais artificiais (*Artificial Neural Networks* - *ANN*).

### 2.2.5 Modelos do Tipo Caixa Cinzenta

Na maioria das vezes, os conhecimentos adquiridos acerca de um dado sistema são escassos e provenientes de diversas fontes. A escassez de conhecimentos fenomenológicos ou de conhecimentos especializados ou ainda a falta de detalhes no conjunto de

dados de entrada e saída de um dado sistema levam à construção de modelos baseados na combinação dos vários conhecimentos existentes e disponíveis, como é o caso da modelização do tipo caixa cinzenta.

Esta designação refere-se a modelos que possuem um razoável nível de interpretabilidade. Estes modelos são construídos com base em conhecimentos fenomenológicos e/ou experimentais cujos respectivos parâmetros internos são ajustados com base em dados de entrada e saída. A interpretabilidade destes modelos depende da proximidade e da semelhança destes, aos modelos do tipo caixa branca ou aos modelos do tipo caixa negra (figura 2.5 [Lit02]). Esta interpretabilidade aumenta com o aumento da quantidade de conhecimentos fenomenológicos utilizados na sua construção.

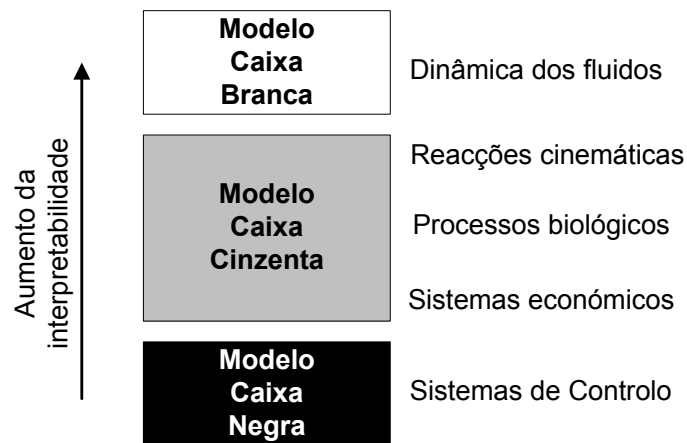


Figura 2.5: Interpretabilidade fenomenológica nos vários tipos de modelos.

Os diferentes tipos de conhecimentos existentes e disponíveis sobre um dado sistema podem-se complementar ou sobrepor. Isto significa que existem sistemas, ou partes destes, que podem ser simultaneamente modelizados com diferentes níveis de interpretabilidade.

Na modelização cinzenta, a definição da contribuição (ou do peso) de cada um dos tipos de conhecimentos é muitíssimo importante, uma vez que a interpretabilidade do modelo depende dos conhecimentos disponíveis sobre o sistema e do modo como estes contribuem e se combinam para a construção do modelo final [Oli98] e [KL96].

Os modelos devem ser construídos com base nos conhecimentos que melhor representam local, ou globalmente, o sistema. Se existir uma determinada região de funcionamento do sistema, em que a melhor, ou única, fonte de conhecimentos sejam os dados de entrada e saída, essa região deve ser modelizada à custa deste tipo de conhecimento. O mesmo poderá acontecer com outras regiões em que estejam disponíveis e utilizados outros tipos de conhecimentos. Deste modo, o modelo final pode ser constituído por vários sub-modelos que, associados em série ou em paralelo, vão contribuir para a construção do modelo global final (figura 2.6).

As várias formas de combinar os diferentes tipos de conhecimentos dão origem a várias metodologias de modelização cinzenta, desde modelos do tipo caixa cinza claro (baseados em conhecimentos fenomenológicos) até aos modelos do tipo caixa cinza

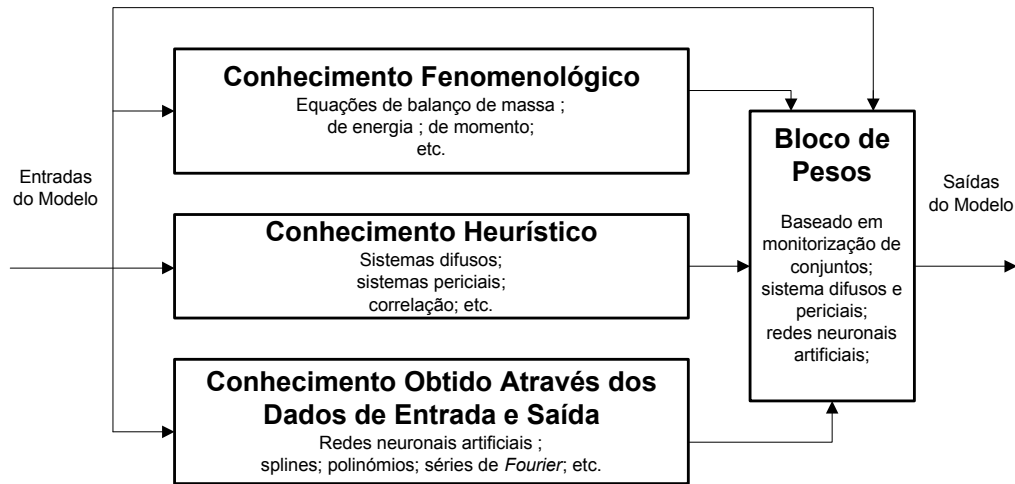


Figura 2.6: Modelo cinzento genérico (combinação dos vários tipos de conhecimento na construção do modelo final).

escuro (baseados em dados de entrada e saída).

O estudo realizado neste trabalho centra-se nos métodos de modelização do tipo caixa cinzenta, mais particularmente, na modelização neuro-difusa, modelização orientada por blocos e modelização híbrida. Estes três tipos de modelização cinzenta possuem diferentes níveis de interpretabilidade o que levou, como se ilustra a figura 2.7, ao seu enquadramento dentro dos modelos do tipo caixa cinzenta escura, cinzenta e cinzenta clara, respectivamente.

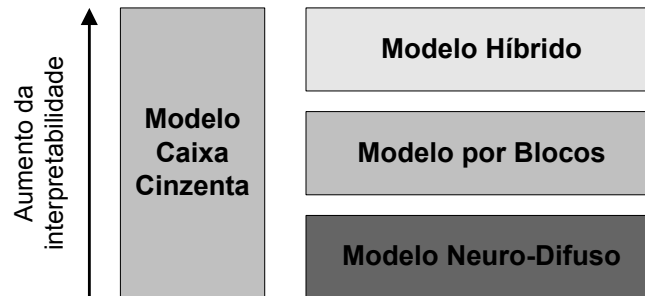


Figura 2.7: Interpretabilidade dos diferentes modelos do tipo caixa cinzenta: modelo híbrido, modelo orientados por blocos e modelo neuro-difuso.

Em primeiro lugar, são abordados e apresentados os modelos neuro-difusos (modelos do tipo caixa cinzenta escura) que utilizam na sua construção conhecimentos heurísticos ou experimentais qualitativos e dados de entrada e saída; em segundo, os modelos orientados por blocos (modelos do tipo caixa cinzenta) que utilizam na sua construção conhecimentos experimentais e/ou fenomenológicos e dados de entrada e saída; finalmente, são abordados e apresentados os modelos híbridos (modelos do tipo caixa cinzenta clara) que na sua construção utilizam essencialmente conhecimentos

fenomenológicos e dados de entrada e saída.

### Modelos Neuro-Difusos

Os modelos designados por modelos neuro-difusos surgem da aplicação dos sistemas neuro-difusos à modelização de sistemas. Os sistemas neuro-difusos são estruturas matemáticas que permitem aproximar qualquer função linear ou não linear com um determinado nível de precisão [BH91], [LL91], [WM92] e [JS93]. Estes sistemas baseiam-se nos sistemas difusos [Zad65], [Zad73], nos quais o conhecimento heurístico ou experimental é expresso através de um conjunto de regras que descrevem o comportamento de um dado sistema ou função e cujos parâmetros internos são ajustados utilizando métodos de aprendizagem típicos das redes neuronais.

Os sistemas difusos, como se ilustra na figura 2.8, são constituídos por quatro blocos fundamentais: a fusificação<sup>2</sup>, a inferência difusa, a base de regras e, por fim, a desfusão<sup>3</sup>.

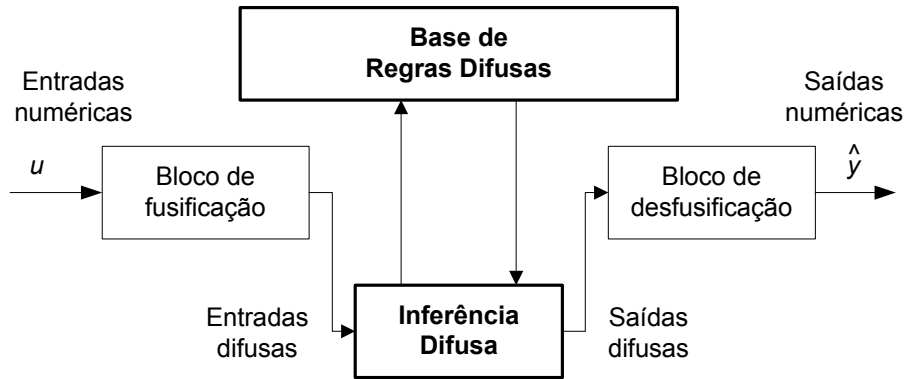


Figura 2.8: Sistema difuso base da modelização neuro-difusa.

A fusificação transforma os dados numéricos em variáveis linguísticas representadas por conjuntos difusos e seus respectivos graus de pertença.

Para melhor compreensão da forma como os sistemas difusos descrevem um dado sistema apresenta-se a descrição difusa de um exemplo muito simples. O exemplo consiste na modelização de um sistema de aquecimento através da relação estática entre a potência de aquecimento e a temperatura de saída. A descrição deste sistema pode ser efectuada através do seguinte conjunto de regras linguísticas:

- SE a potência de aquecimento é reduzida, ENTÃO a temperatura é baixa;
- SE a potência de aquecimento é média, ENTÃO a temperatura é moderada;
- SE ...

<sup>2</sup>Fusificação - tradução do inglês *fuzzification*. Esta tradução não é a única (podem aparecer outras como fuzificação ou difusificação).

<sup>3</sup>Desfusificação - tradução do inglês *defuzzification*. Esta tradução não é a única (podem aparecer outras como desfuzificação ou desdifusificação).

Estas regras traduzem conhecimento experimental e qualitativo sobre os sistemas à custa da definição de conjuntos difusos (por exemplo: *forte*, ..., *fraca*, *alta*, ..., *baixa*), transformando os conhecimentos qualitativos em relações quantitativas. A aplicação dos métodos de inferência ao conjunto de regras permite obter uma conclusão qualitativa relativamente às entradas do sistema.

A desfusão é precisamente o processo inverso à fusificação, isto é, transforma a conclusão difusa qualitativa numa conclusão numérica quantitativa.

Quando os sistemas difusos são expressos através de funções matemáticas semelhantes às utilizadas na definição das redes neuronais, as técnicas de aprendizagem típicas destas podem ser utilizadas nos sistemas difusos. Assim, surgem os sistemas neuro-difusos que são sistemas difusos que utilizam técnicas de aprendizagem baseadas na derivada do erro permitindo o ajuste da respectiva estrutura interna.

Existem muitas aplicações bem sucedidas destas metodologias descritas na literatura provando a sua capacidade universal de aproximação de funções [LL91], [BK92], [Nau94], [Jan93] e [LC99]. Três resumos sobre as várias abordagens de modelização e identificação neuro-difusa que surgiram na última década podem ser encontrados em [Abr01], [AN00] e [VM04d].

A construção deste tipo de modelos utiliza conhecimentos heurísticos ou experimentais criando modelos com algum nível de interpretabilidade. A utilização dos dados de entrada e saída e os critérios de identificação utilizados podem diminuir fortemente a interpretabilidade destes modelos - daí a classificação destes como modelos do tipo caixa cinza escuro.

Os sistemas neuro-difusos (em particular a sua aplicação na modelização neuro-difusa) serão abordados em detalhe no capítulo 3.

### Modelos Orientados por Blocos

O método de modelização orientado por blocos baseia-se, como ilustra a figura 2.9, na combinação de dois blocos: a ligação em série de uma função estática  $f(\cdot)$  não linear seguida de um modelo dinâmico linear  $G(\cdot)$  conhecida por modelo Hammerstein; ou a ligação em série de um modelo linear  $G(\cdot)$  seguido de uma função estática não linear  $f(\cdot)$  conhecida por modelo Wiener.

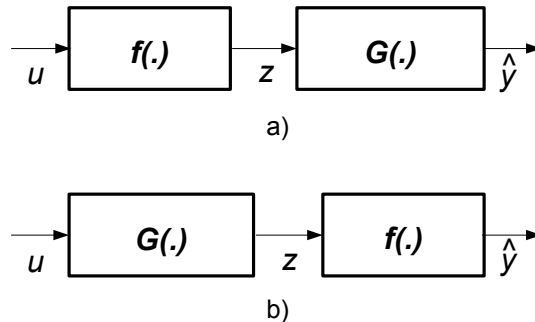


Figura 2.9: Modelização orientada por blocos - a) modelo Hammerstein b) modelo Wiener.

Esta modelização permite a construção de modelos dinâmicos não lineares para uma determinada classe de sistemas. Esta classe é caracterizada por sistemas que apresentam um comportamento linear dinâmico e uma não linearidade estática. Existem várias aplicações na literatura que comprovam que modelos orientados por blocos de sistemas não lineares pouco complexos fornecem modelos muito mais precisos do que os fornecidos pelos modelos lineares clássicos [ZS94], [EJL91b], [NPR96] e [MAP97]. A construção do bloco dinâmico linear é baseada nas técnicas de modelização lineares clássicas que utilizam os dados de entrada e saída para a estimação dos respectivos parâmetros. Para a construção do bloco estático não linear podem ser usadas várias técnicas como, por exemplo, redes neuronais, sistemas difusos, aproximação polinomial, equações fenomenológicas, etc.

A divisão do modelo em dois blocos e a escolha do tipo de modelo orientado por blocos a utilizar é baseada em conhecimentos fenomenológicos e/ou conhecimentos experimentais adquiridos *a priori*.

Este tipo de modelos é construído à custa de conhecimentos fenomenológicos, heurísticos e conhecimentos obtidos através dos dados de entrada e saída do sistema. Contudo, a utilização de conhecimentos fenomenológicos e heurísticos não é muito elevada, gerando modelos com um grau de interpretabilidade não muito elevado.

Como a construção e identificação dos blocos linear (e por vezes também do bloco não linear) é feita à custa de dados de entrada e saída, estes modelos podem possuir partes não interpretáveis - daí a sua classificação como modelos do tipo caixa cinzenta.

No capítulo 4 serão abordados, descritos e caracterizados com maior detalhe os métodos de modelização orientada por blocos Hammerstein e Wiener.

## Modelos Híbridos

O método de modelização híbrido é baseado, como ilustra a figura 2.10, na ligação em série ou em paralelo, de um modelo fenomenológico do sistema e de um modelo do tipo caixa cinzenta ou negra. O modelo do tipo caixa cinzenta ou negra, a ligar em série ou paralelo com o modelo fenomenológico, modeliza a parte do modelo fenomenológico que é mal conhecida ou até desconhecida.

Neste método, a maior componente do modelo é construída através dos conhecimentos fenomenológicos do sistema (equações de balanço de massa, de energia, de momento, etc.).

Existem várias aplicações na literatura que comprovam que modelos híbridos são modelos relativamente simples, com elevada capacidade de generalização e com um elevado nível de interpretabilidade [PU92b], [TK94b], [Oli98], [Lit02] e [XJ02].

Na construção deste tipo de modelos, os conhecimentos fenomenológicos aplicados são determinantes, gerando modelos quase totalmente interpretáveis. No entanto, a utilização dos dados de entrada e saída para a identificação das partes desconhecidas do sistema pode contribuir para que estes modelos contenham algumas partes de menor interpretabilidade - daí a classificação destes como modelos do tipo caixa cinza claro.

Os vários métodos de modelização híbrida vão ser abordados e explorados com mais detalhe no capítulo 5.

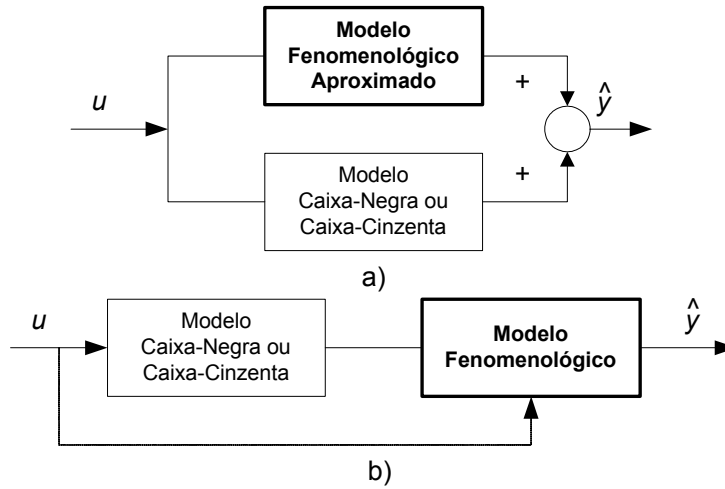


Figura 2.10: Modelização híbrida a) modelo paralelo b) modelo série.

## 2.3 Identificação de Sistemas

A identificação de sistemas é o processo de estimação/ajuste de um conjunto de parâmetros desconhecidos de uma dada estrutura matemática discreta (modelo do sistema) de forma a que a resposta do modelo seja idêntica à resposta do sistema.

A escolha da estrutura matemática utilizada para a construção do modelo, como as estruturas neuro-difusas, orientadas por blocos e híbridas, tem um papel fundamental no processo de identificação de sistemas e nas características do modelo final identificado; no entanto, o processo de identificação não se limita à escolha da estrutura matemática (método de modelização), realizando ainda outras tarefas não menos importantes, como a selecção da condição experimental, a aquisição de dados e a selecção dos critérios de identificação e validação.

### 2.3.1 Ciclo de Identificação

A identificação de sistemas consiste, genericamente na definição da estrutura do modelo e no ajuste dos seus parâmetros tendo por base dados experimentais de entrada e saída do sistema e conhecimentos *a priori*. Mais especificamente, poder-se-á dizer que a identificação é uma tarefa influenciada por cinco factores fundamentais [SS89] e [Lju87]:

- o sistema ( $S$ );
- uma condição experimental ( $C$ ) para a aquisição de dados experimentais;
- um modelo/estrutura ( $M$ ) que represente o sistema;
- um critério de identificação ( $I$ ) que estime os parâmetros do modelo;
- um critério de validação ( $V$ ) que verifique o erro de modelização e a sua capacidade de generalização.



Para a identificação de sistemas discretos é necessário um conjunto de dados  $\mathbf{D}^{ND}$ , com  $ND$  amostras de dados de entrada ( $u(k)$ ) e de saída ( $y(k)$ ) do sistema (ver equação 2.1), conforme definidos na condição experimental,  $C$  (para simplicidade de notação, a equação 2.1 representa os dados de entrada e saída de um sistema SISO):

$$\mathbf{D}^{ND} = \{u(k), y(k)\}_{k=1}^{ND} \quad (2.1)$$

A identificação de sistemas implica a definição de uma estrutura matemática  $M$ . Conforme descrito na secção 2.2, existem vários métodos para a definição da estrutura matemática do modelo  $M$ , incluindo modelos fenomenológicos, modelos polinomiais, modelos lineares clássicos, modelos difusos, modelos neuro-difusos, redes neuronais, etc.

Normalmente, os modelos são restritos a uma família de funções definidas num conjunto finito de amostras passadas de entradas e saídas. Estas famílias são parametrizadas através de um vector  $\theta$ , permitindo a definição de uma aproximação da função do sistema (como definido na equação 2.2):

$$\hat{y} = g(\mathbf{D}^{ND}; \theta) \quad (2.2)$$

em que  $\mathbf{D}^{ND}$  é o conjunto limitado de  $ND$  amostras dos sinais de entrada e saída,  $\theta$  o vector dos parâmetros a determinar, e  $\hat{y}$  a saída estimada do modelo.

Esta forma de modelizar é conhecida por modelização paramétrica [Lju87], uma vez que o modelo é construído à custa de um conjunto de dados discretos de entrada e saída  $\mathbf{D}^{ND}$  e do vector de parâmetros a determinar  $\theta$ .

O objectivo da identificação é determinar o vector dos parâmetros  $\theta$  que melhor aproxima a saída do modelo  $\hat{y}$  à saída do sistema real  $y$ . A determinação do vector  $\theta$  é realizada com base no vector de erro  $e(k)|_{k=1..ND}$ , vector que contém as diferenças entre as saídas do modelo e as saídas do sistema real, como ilustra a figura 2.11.

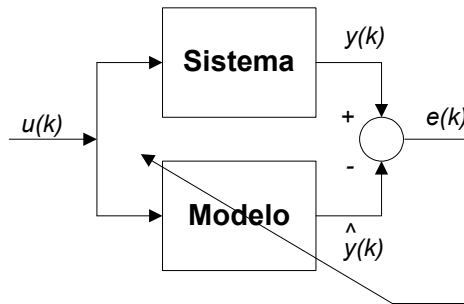


Figura 2.11: Identificação do modelo do sistema através do sinal  $e(k)$ .

Para um dado conjunto de dados de entrada e saída do sistema, o processo de minimização do vector de erro  $e(k)|_{k=1..ND}$ , expresso na equação 2.3, implica a utilização de uma função de custo, designada por  $J$ , que pode ser definida através de várias expressões matemáticas. Um exemplo muito utilizado de uma função de custo, para a identificação fora de linha, é o erro quadrático médio expresso pela equação 2.4:

$$e(k) = [y(k) - \hat{y}(k)]|_{k=1..ND} \quad (2.3)$$

$$J = \frac{1}{ND} \sum_{k=1}^{ND} e(k)e(k)^T \quad (2.4)$$

Com a definição da função de custo  $J$  e do método de identificação utilizado (critério de identificação  $I$ ), o vector dos parâmetros  $\theta$  é calculado.

Por fim, é definido o critério de validação  $V$  e o modelo é avaliado.

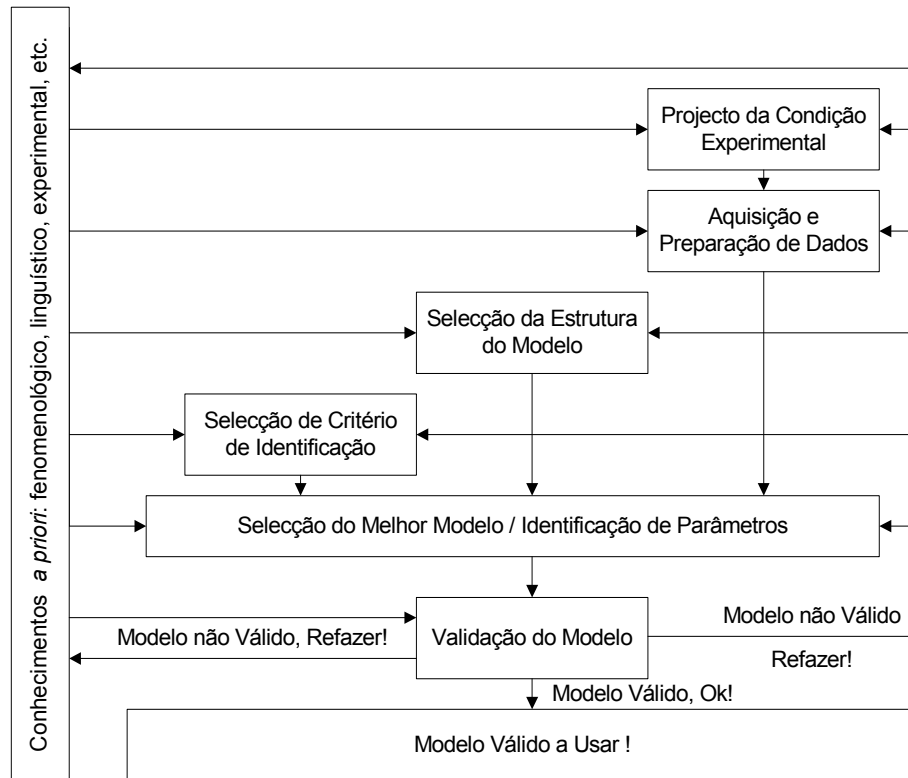


Figura 2.12: Ciclo de identificação.

Assim, de acordo com a figura 2.12, a identificação de sistemas é um processo iterativo que só termina quando são atingidos os pressupostos definidos nas várias etapas do ciclo de identificação, mais particularmente, nos critérios de identificação  $I$  e validação  $V$ .

### 2.3.2 O Sistema

O sistema a modelizar condiciona fortemente o tipo de modelo a utilizar e este, por sua vez, condiciona o erro de modelização, a flexibilidade e a interpretabilidade final do modelo construído.

Na realidade todos os sistemas são, em última instância, não lineares. No entanto, as não linearidades existentes em alguns sistemas são pouco significativas, permitindo que a utilização de modelos lineares (em torno de um ponto de operação) seja satisfatória. Quando isso não acontece, é essencial utilizar um método de modelização não linear para melhor representar o comportamento do sistema.

Os algoritmos mais utilizados na detecção de não linearidades são [Hab85]:

- teste no domínio do tempo;
- teste em regime estacionário;
- método de correlação cruzada não linear;
- método de autocorrelação de ordem elevada.

Estes algoritmos permitem verificar a existência de não linearidades, que tipo de não linearidades existem (estáticas ou dinâmicas) e entre que variáveis de entrada e saída é que elas se encontram. A linearidade ou não linearidade e o conhecimento *a priori* das características do sistema em causa condicionam o tipo de modelo a utilizar e as características do modelo final construído.

### 2.3.3 Condição Experimental

A primeira etapa para a realização do ciclo de identificação é a definição da condição experimental. Esta inclui a definição das variáveis significativas do sistema, a definição do período de amostragem, a recolha de dados de entrada e saída, o número de amostras e o pré-processamento dos dados.

A definição da condição experimental é responsável pela quantidade e qualidade dos dados adquiridos, os quais devem conter o máximo de riqueza informativa possível. A correcta definição da condição experimental está revestida de grande importância uma vez que a recolha dos dados pode ser economicamente dispendiosa e demorada. Esta recolha implica testes reais no sistema que por vezes são demorados e até difíceis de realizar, especialmente, em processos industriais já instalados.

A definição do período de amostragem deve ser tal que permita a correcta aquisição dos sinais para que os aspectos fundamentais da sua dinâmica e do seu estado estacionário sejam retidos. Para sistemas lineares existem alguns critérios, bastante conhecidos, que orientam a selecção do intervalo de amostragem [AW84] e [Lju87]. Tipicamente, as propostas encontradas baseiam-se numa relação entre frequência de amostragem e largura de banda do sistema em malha aberta. Assim, a escolha do intervalo de amostragem pode basear-se na largura de banda do sistema que pode ser equivalente ao seu tempo de subida [Mot97].

Uma forma prática que estabelece uma relação entre o tempo de subida e o tempo de amostragem, para sistemas de primeira ordem, é a definida na equação 2.5 [AW84]:

$$h = \frac{\tau_s}{N_r} \quad \text{em que } N_r \simeq 4 \text{ a } 10 \quad (2.5)$$

em que  $\tau_s$  é o tempo de subida,  $h$  é o período de amostragem e  $N_r$  é o número de amostras que o tempo de subida deve conter.

Outro método é sugerido em [Lju87], definindo a frequência de amostragem dez vezes superior à largura de banda do sistema em malha fechada.

Para sistemas não lineares não existem regras genéricas para a definição do período de amostragem como as regras existentes para sistemas lineares; no entanto, é típico utilizarem-se as mesmas orientações.

Os sinais de entrada aplicados ao sistema para a aquisição dos dados de identificação devem percorrer todos os valores possíveis e ter uma elevada persistência de excitação. Quanto maior for a persistência de excitação dos sinais de entrada, mais rica é a qualidade dos dados obtidos na saída. Mais uma vez, existem regras bem conhecidas para a definição dos sinais de entrada a aplicar nos sistemas lineares que garantem a convergência dos parâmetros dos respectivos modelos lineares. Por outro lado, para sistemas não lineares só existem regras qualitativas para a definição dos sinais de entrada que, simplesmente, potenciam uma boa modelização.

Para este tipo de sistemas, é necessário escolher um sinal de entrada não só com um espectro de frequências alargado, como também uma gama de amplitudes abrangentes [Sor94]. O procedimento habitual consiste em utilizar, como sinais excitadores da entrada do sistema, a soma de muitos sinais sinusoidais de várias amplitudes e frequências ou sinais aleatórios.

Por outro lado, é necessário decidir se os dados são recolhidos em malha aberta ou fechada. Tipicamente, são utilizados dados em malha aberta que se mantenham dentro dos valores onde o sistema é suposto operar. A utilização de dados em malha fechada pode levar à perda de capacidade de identificação, devido à possível existência de correlação entre os sinais de entrada e saída do sistema [SS89].

Finalmente, é típico efectuar um pré-processamento dos dados recolhidos, especialmente se as medidas efectuadas apresentarem níveis de ruído significativos relativamente à amplitude dos próprios sinais. Este pré-processamento baseia-se fundamentalmente numa filtragem das altas frequências. Os filtros passa-baixo utilizados devem ser escolhidos por forma a que não mascarem as características intrínsecas dos sinais. Outro método utilizado para redução do efeito do ruído é o aumento, quando possível do período de amostragem [SS89]. Deste modo, a covariância do ruído existente nas medidas dos sinais diminui de magnitude, enfatizando as propriedades de baixa frequência dos sinais do sistema.

### 2.3.4 Selecção da Estrutura do Modelo

A selecção da estrutura matemática ou do tipo de modelo a utilizar na identificação de sistemas é uma das decisões mais importantes e mais complexas em todo o processo de identificação.

A estrutura dos modelos deve ser escolhida com base nos conhecimentos *a priori* e norteada segundo um compromisso de flexibilidade, simplicidade e interpretabilidade da classe de modelos considerada. A escolha de uma estrutura demasiado simples, com reduzido número de parâmetros internos, isto é subparametrizada, poderá reflectir-se na incapacidade de representação do sistema. Por outro lado, a escolha de uma estrutura

demasiado complexa, com elevado número de parâmetros internos, ou seja sobreparametrizada, poderá reflectir-se num sobreajustamento dos parâmetros aos dados e num aumento desnecessário da complexidade da estrutura.

O número óptimo de parâmetros para um dado modelo será aquele que respeite o compromisso entre flexibilidade e parcimónia de modelos. Modelos com maior flexibilidade implicam estruturas com maior capacidade de modelização. Isto só é conseguido à custa do aumento do número de parâmetros. A parcimónia de um modelo implica uma reduzida complexidade matemática da estrutura, o que se consegue à custa de uma utilização moderada do número de parâmetros. Defende este princípio que não se devem utilizar parâmetros em excesso se não forem estritamente necessários. O tipo de estrutura matemática utilizada influencia a qualidade do modelo final e o custo a pagar por ele.

Perante estes problemas, torna-se clara a importância da escolha adequada da estrutura matemática do modelo. Esta estrutura depende essencialmente de factores como a classe ou o tipo de modelo, a dimensão do modelo e a sua parametrização [Lju87].

### Tipo de Modelo

Como regra geral, a primeira estrutura do modelo a utilizar, deve ser uma estrutura simples. Mesmo perante sistemas aparentemente complexos ou não lineares, a primeira estrutura a utilizar deve ser uma estrutura linear, visto que estas estruturas possuem técnicas simples e robustas de estimação dos respectivos parâmetros internos gerando estruturas de baixa complexidade. Só devem ser utilizados métodos de modelização não lineares mais complexos quando os métodos de modelização lineares mais simples não verificarem os critérios de identificação e validação predefinidos.

No entanto, se o sistema for efectivamente não linear, existe a possibilidade de incluir essas não linearidades na estrutura matemática, mantendo os seus parâmetros lineares. Por exemplo, a modelização *semi-física* [Lin96] que modeliza as não linearidades entre as variáveis do sistema - relações quadráticas, logarítmicas, etc. - através das respectivas leis fenomenológicas, mantendo lineares, os parâmetros a identificar. Outro exemplo, é a modelização das não linearidades introduzidas pelas funções não lineares conhecidas dos sensores e/ou actuadores nas grandezas medidas e/ou de actuação [KR76].

No caso em que as não linearidades são totalmente desconhecidas, utilizam-se métodos de modelização não lineares. Estes implicam estruturas matemáticas complexas e algoritmos de optimização não lineares, que influenciam a complexidade do algoritmo de aprendizagem e a interpretabilidade do modelo.

### Dimensão do Modelo

A dimensão do modelo depende, fundamentalmente, da selecção dos regressores a utilizar (ordem do modelo, atrasos associados a cada variável) e da selecção das variáveis fenomenológicas a incluir.

A escolha dos regressores é efectuada à custa dos dados de entrada e saída e de conhecimentos adquiridos *a priori*. Para a construção e identificação do modelo devem

ser utilizadas todas as variáveis controláveis ou não controláveis conhecidas e mensuráveis que actuam no sistema. Por outro lado, quanto maior é o número de entradas do modelo (número de regressores), maior é o número de parâmetros e a complexidade matemática da estrutura utilizada.

Um dos princípios básicos a ser utilizado na escolha dos regressores é o princípio da parcimónia. Este princípio, quando aplicado à modelização de sistemas, defende que: “não se devem usar parâmetros extra para a descrição de fenómenos dinâmicos se estes não forem, estritamente, necessários” [Lju87]. Segundo o mesmo, dos vários modelos que representam correctamente o sistema real, o modelo com o menor número de parâmetros independentes ou menor complexidade matemática deverá ser o escolhido.

Para sistemas lineares existem métodos baseados na análise dos dados de entrada e saída que ajudam a resolver os problemas da sobreparametrização, como o teste das matrizes de co-variância; correlação de variáveis, exame da matriz de informação. Para a definição da ordem de sistemas existe, por exemplo, o critério de Akaike (*Akaike's Information Criterion - AIC*) [Aka73] muito conhecido. A aplicação dos métodos lineares na escolha dos regressores dos modelos não lineares pode ser efectuada, contudo, não existe a garantia de que os regressores determinados sejam os mais adequados.

A escolha dos regressores em modelos não lineares deve basear-se nos conhecimentos fenomenológicos existentes sobre o sistema. Na ausência destes, pode utilizar-se, por exemplo, o método da força bruta, que consiste no teste sequencial das várias hipóteses de regressores, seleccionando os regressores que obtiverem resultados de modelização satisfatórios para uma determinada classe ou tipo de modelos evitando a construção de modelos sobreparametrizados.

## Parametrização do Modelo

Os métodos de modelização estudados no âmbito deste trabalho utilizam modelos com estruturas matemáticas parametrizadas. A introdução da dinâmica neste tipo de modelos é efectuada utilizando a estrutura série/paralelo, ilustrada na figura 2.13. Esta estrutura inclui a dinâmica nos modelos paramétricos introduzindo linhas de atraso, ou seja, introduzindo entradas e saídas passadas como entradas do modelo paramétrico. Assim, a introdução destas linhas de atraso permite a construção de modelos com dinâmica capazes de aproximar sistemas dinâmicos reais.

A figura 2.13 apresenta a estrutura série/paralela para um sistema SISO. Nesta figura  $u(k)$  e  $y(k)$  representam a entrada e a saída do sistema,  $\hat{y}(k)$  representa a saída do modelo (estimação da saída do sistema),  $n_u$  é o número de regressores da variável de entrada  $u$ ,  $n_y$  é o número de regressores da variável de saída  $y$ ,  $d$  é o tempo morto entre a entrada e saída do sistema ( $d$  número inteiro múltiplo de  $h$ , sendo  $h$  o período de amostragem) e Atraso representa o atraso de uma amostra.

A estimação da saída  $\hat{y}(k)$  do modelo paramétrico genérico pode ser obtida através da equação 2.6:

$$\hat{y}(k) = g(\varphi(k); \theta) \quad (2.6)$$

em que  $\varphi(k)$  é o vector dos regressores do modelo, expresso na equação 2.7 e  $\theta$  o vector dos parâmetros do modelo a determinar.

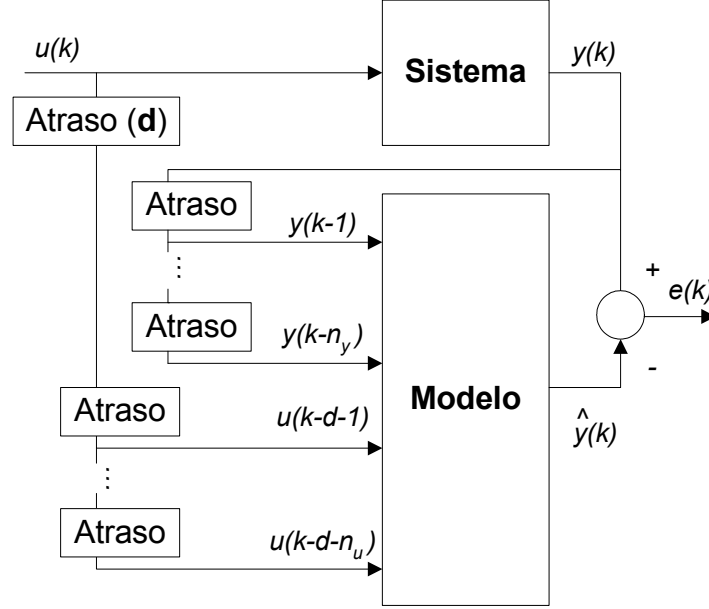


Figura 2.13: Estrutura paramétrica série/paralelo.

$$\boldsymbol{\varphi}(k) = [y(k-1) \dots y(k-n_y) \ u(k-d-1) \dots u(k-d-n_u)]^T \quad (2.7)$$

Depois da escolha dos regressores e dos respectivos atrasos, da selecção das variáveis fenomenológicas a incluir e da definição da estrutura matemática do modelo, há que estimar ou identificar o vector dos parâmetros  $\boldsymbol{\theta}$  do respectivo modelo minimizando uma função de custo baseada no erro  $e(k)$  definido na equação 2.8:

$$e(k) = y(k) - g(\boldsymbol{\varphi}(k); \boldsymbol{\theta}) \quad (2.8)$$

### 2.3.5 Critérios de identificação

Os critérios de identificação dependem da função de custo a minimizar e do método de identificação utilizado. O objectivo da identificação é calcular ou ajustar o vector  $\boldsymbol{\theta}$ , para minimizar uma função de custo  $J(\boldsymbol{\theta}, \mathbf{D}^{ND})$ . A minimização desta função depende, normalmente, do vector de erros  $e(k)|_{k=1..N}$ , dos dados  $\mathbf{D}^{ND}$  (dados de treino), do vector dos parâmetros  $\boldsymbol{\theta}$  e da função de custo  $J$ .

A função de custo é normalmente definida como o somatório do erro quadrático médio (*Mean Square Error* - *MSE*), expresso na equação 2.9, ou a raiz quadrada do erro quadrático médio (*Root Mean Square Error* - *RMSE*), definido na equação 2.10:

$$J(\boldsymbol{\theta}, \mathbf{D}^{ND}) = \frac{1}{ND} \sum_{t=1}^{ND} e(k)e(k)^T \quad (2.9)$$

$$J(\boldsymbol{\theta}, \mathbf{D}^{ND}) = \sqrt{\frac{1}{ND} \sum_{t=1}^{ND} e(k)e(k)^T} \quad (2.10)$$

O cálculo do vector dos parâmetros  $\boldsymbol{\theta}$  é obtido minimizando o argumento da função de custo  $J(\boldsymbol{\theta}, \mathbf{D}^{ND})$ , como se define na equação 2.11:

$$\boldsymbol{\theta} = \boldsymbol{\theta}(\mathbf{D}^{ND}) = \min \arg [J(\boldsymbol{\theta}, \mathbf{D}^{ND})] \quad (2.11)$$

A função de custo utilizada pode ser função de outras variáveis além do erro como, por exemplo, o número de parâmetros do modelo. Quando se pretende ter em conta a complexidade do modelo, uma das formas possíveis de o conseguir é incorporar o número de parâmetros na função utilizada, de modo a obter um compromisso entre o erro de modelização e a complexidade do modelo.

Além da definição da função de custo  $J(\boldsymbol{\theta}, \mathbf{D}^{ND})$  a minimização desta depende do método de identificação utilizado. A escolha do método de identificação depende do tipo de parâmetros lineares ou não lineares e da própria estrutura matemática do modelo.

Para a identificação de parâmetros lineares existem métodos de identificação solidamente estudados que garantem a convergência dos parâmetros para o mínimo global como, por exemplo, o método de identificação dos mínimos quadrados (*Least Squares - LS*) para sistemas não variantes no tempo e o método recursivo dos mínimos quadrados (*Recursive Least Squares - RLS*) para sistemas variantes no tempo. Estes métodos apresentam algoritmos computacionalmente simples que garantem convergência dos parâmetros minimizando globalmente uma dada função de custo.

Para a identificação de parâmetros não lineares não existe uma teoria solidamente estudada que garanta a convergência dos parâmetros para o mínimo global. Existem métodos que só garantem a convergência para o mínimo global, mediante a verificação de um conjunto de restrições que nem sempre são possíveis de garantir. Os métodos de identificação mais conhecidos para a estimação de parâmetros não lineares são os métodos baseados no gradiente (*Gradient Descent - GD*) e os métodos baseados nos algoritmos genéticos (*Genetic Algorithm - GA*).

Os métodos baseados no gradiente baseiam-se, como o próprio nome indica, na derivada da função de custo em ordem aos parâmetros a identificar. O objectivo do algoritmo é corrigir progressivamente os parâmetros de modo a minimizar a função de custo.

Deste método existem duas variantes muito conhecidas, o método da *Retropropagação do erro*, que diminui uma função de custo utilizando as derivadas de primeira ordem, e os métodos de *Newton* e *Levenberg-Marquardt* (*Levenberg-Marquardt - LM*) [Lev44], [Mar63], que diminuem a função de custo utilizando as derivadas de segunda ordem. Os métodos que utilizam as derivadas de segunda ordem convergem mais rapidamente, no entanto, possuem algoritmos de aprendizagem bastante mais complexos.

Estes métodos podem convergir para um mínimo local. Para evitar esta situação é necessário utilizar técnicas que permitam a estes algoritmos sair do mínimo local e convergir para o mínimo global, como a técnica da variação do factor de aprendizagem [JSM97].



Os métodos baseados nos algoritmos genéticos utilizam algoritmos de evolução genética (evolução de “pais” para “filhos”). Inicialmente, testa-se um grupo de soluções aleatórias e deste grupo é guardado um subgrupo, “os pais”, que melhor minimizem a função de custo. De seguida, desse grupo são gerados “os filhos”, parâmetros com alguma relação ao do subgrupo “pais”, e testam-se novamente estas soluções guardando desse subgrupo as soluções que minimizaram a função de custo. Este algoritmo é executado enquanto ocorrerem soluções que minimizem cada vez mais a referida função de custo e/ou enquanto a função de custo não atinja um determinado valor predefinido.

### 2.3.6 Critérios de Validação

A avaliação da qualidade dos modelos é o último passo do ciclo de identificação e consiste na sua validação numérica, observando o erro de modelização e a capacidade de generalização. Esta avaliação é feita com base num conjunto de dados de entrada e saída, conhecidos por dados de teste  $\mathbf{H}^{NH}$  (equação 2.12), e no vector de parâmetros  $\theta$  identificado na fase de identificação (fase de treino):

$$\mathbf{H}^{NH} = \{u(k), y(k)\}_{k=1}^{NH} \quad (2.12)$$

O conjunto de dados de teste  $\mathbf{H}^{NH}$  é em tudo semelhante ao conjunto de dados de treino  $\mathbf{D}^{ND}$  (sinais com variações em amplitude e frequência idênticas) em que  $NH$  é o número de amostras.

O critério de validação permite a comparação da saída dos modelos com a saída do sistema, determinando se o modelo obtido representa suficientemente bem o sistema. Este procedimento pode ser efectuado utilizando vários critérios [Lit02]:

- cruzamento de parâmetros;
- erro;
- erro quadrático médio;
- raiz quadrada do erro quadrático médio (em alternativa ao anterior).

O critério de validação por cruzamento de parâmetros consiste na utilização de dois vectores de dados de entrada e saída semelhantes para a identificação dos parâmetros de um dado modelo. Se os parâmetros obtidos não diferirem muito uns dos outros, o modelo pode ser considerado de boa qualidade; caso contrário, os modelos têm de ser revistos.

O critério de validação baseado no erro consiste na observação do vector do sinal de erro  $e(k)|_{k=1..NH}$ , também chamado sinal inovação [Lit02]. Este critério fornece informação sobre a diferença existente entre as saídas do modelo e do sistema, bem como das partes não modelizadas do sistema. Para um determinado modelo ideal este sinal de erro deve obedecer a uma distribuição normal com média nula e variância baixa. Se não respeitar a referida distribuição, significa que o modelo não representa correctamente o sistema.

O critério de validação baseado no erro quadrático médio  $V(\boldsymbol{\theta}, \mathbf{H}^{NH})$ , é semelhante ao critério de identificação  $J(\boldsymbol{\theta}, \mathbf{D}^{ND})$  (equação 2.9), só que aplicando os dados de entrada e saída de teste  $\mathbf{H}^{NH}$ , e o vector dos parâmetros  $\boldsymbol{\theta}$  identificado, equação 2.13:

$$V(\boldsymbol{\theta}, \mathbf{H}^{NH}) = \frac{1}{NH} \sum_{k=1}^{NH} e(k)e(k)^T \quad (2.13)$$

Este critério é bastante utilizado e permite avaliar o erro de modelização e a capacidade de generalização de um dado modelo.

E por fim, o critério de validação baseado na raiz quadrada do erro quadrático médio, semelhante ao critério de validação anterior, analisa o somatório da raiz quadrada dos erros quadráticos médios, como definido na equação 2.14:

$$V(\boldsymbol{\theta}, \mathbf{H}^{NH}) = \sqrt{\frac{1}{NH} \sum_{k=1}^{NH} e(k)e(k)^T} \quad (2.14)$$

O seu valor permite avaliar o erro de modelização, a capacidade de generalização e o erro cometido pelo modelo construído.

## 2.4 Conclusão

Neste capítulo abordaram-se questões relacionadas com a modelização e identificação de sistemas.

O modelo ideal para um sistema é o modelo contínuo ou analítico e interpretável baseado nas leis fenomenológicas do sistema a modelizar. Contudo, para sistemas relativamente complexos e não lineares, as equações do modelo analítico são geralmente difíceis e, muitas vezes, impossíveis de determinar. Deste modo, surgiram diversos métodos de modelização cinzentos que permitem a construção de modelos paramétricos, mais ou menos interpretáveis, utilizando conhecimentos fenomenológicos, experimentais e obtidos através de dados de entrada e saída do sistema.

A escolha do método de modelização tem um papel fundamental nas características do modelo final; no entanto, o processo de identificação de sistemas não se limita à escolha do método ou do tipo de modelo, realizando ainda outras tarefas não menos importantes.

O processo de identificação de sistemas é uma tarefa influenciada por cinco factores fundamentais: o sistema; a condição experimental; o tipo de estrutura do modelo, o critério de identificação e o critério de validação. Cada um destes factores implica tomadas de decisão importantes que influenciam as características do modelo identificado.

A secção 2.2 definiu o que é um modelo de um sistema, para que serve e quais os tipos de conhecimentos utilizados na sua construção. Apresentaram-se depois vários métodos de modelização, desde modelização do tipo caixa branca até modelização do tipo caixa negra, dando particular atenção a três métodos de modelização do tipo caixa cinzenta: neuro-difusa, orientada por blocos e híbrida. Estes tipos de modelos

possuem características específicas que lhes permitem modelizar sistemas com um determinado grau de interpretabilidade conjugando na sua construção vários métodos de modelização e vários tipos de conhecimentos.

A secção 2.3 mostrou o ciclo de identificação de sistemas e as suas várias etapas. Foi descrita a importância das características do sistema no tipo de modelo a utilizar. Foram apresentados os factores mais importantes a ter em conta na aquisição de dados, em particular, quanto à riqueza de informação. Descreveram-se os factores mais importantes que condicionam a escolha da estrutura do modelo, sendo este passo um dos mais importantes do ciclo de identificação. De seguida, foram apresentados os critérios de identificação que dependem das funções de custo a minimizar, das funções de custo utilizadas e dos métodos de identificação utilizados (função da linearidade / não linearidade dos parâmetros a identificar). Finalmente, apontaram-se os critérios de validação numérica de modelos mais comuns que permitem a avaliação da qualidade e capacidade de generalização do modelo.

Após o enquadramento apresentado neste capítulo sobre modelização e identificação de sistemas, os capítulos seguintes apresentarão detalhadamente os três métodos de modelização cinzenta já referidos. Serão definidos e estudados os aspectos fundamentais das diferentes estruturas matemáticas paramétricas e respectivos métodos de identificação, começando com a modelização e identificação neuro-difusa no capítulo 3, a modelização e identificação orientada por blocos no capítulo 4, terminando com a modelização e identificação híbrida, no capítulo 5.



# Capítulo 3

## Modelização e Identificação Neuro-Difusa

### 3.1 Introdução

Os sistemas neuro-difusos, apresentados no capítulo 2, são classificados como um método de modelização do tipo caixa cinzenta escura, uma vez que são construídos com base em conhecimentos experimentais provenientes de especialistas (conhecimento interpretável) e em conhecimentos provenientes dos dados de entrada e saída (conhecimento não interpretável).

Este capítulo apresenta a constituição, caracterização e identificação de vários sistemas neuro-difusos, mostrando a forma como estes incorporam os vários tipos de conhecimentos e como é que estes condicionam a sua interpretabilidade.

Os sistemas neuro-difusos foram introduzidos no final dos anos 80 por Brown [BH91], Lin e Lee [LL91], Wang [WM92] e Jang [JS93]. Desde então, o desenvolvimento e aplicação deste tipo de sistemas cresceu exponencialmente a avaliar pelo número de publicações nesta área.

Estes sistemas são caracterizados por estruturas semelhantes às dos sistemas difusos em que os parâmetros dos conjuntos difusos e as respectivas regras são ajustadas mediante a apresentação de pares de treino de sinais de entrada e saída. Estes sistemas exibem duas formas distintas de comportamento: numa primeira fase comportam-se como redes neuronais, aprendendo alguns parâmetros e características essenciais ao seu funcionamento futuro, utilizando algoritmos baseados na descida do gradiente. Posteriormente, na fase de processamento, comportam-se simplesmente como sistemas difusos.

Este capítulo começa por apresentar os sistemas difusos e as redes neuronais artificiais, técnicas que deram origem aos sistemas neuro-difusos. Apresenta os fundamentos dos sistemas neuro-difusos, os vários tipos de estruturas e respectivos parâmetros. Descreve também várias técnicas de identificação dos vários sistemas neuro-difusos, assim como os processos de inicialização destas estruturas. Por fim, faz um levantamento de vários sistemas neuro-difusos desenvolvidos e apresentados ao longo da última década.

## 3.2 Sistemas Difusos

Os sistemas difusos [DHR93] e [Wan97] derivam da teoria dos conjuntos difusos apresentada na década de 60 por Zadeh [Zad65] e do desenvolvimento da lógica difusa, permitindo a descrição matemática de sistemas por meio de um conjunto de regras qualitativas como as usadas no dia-a-dia através da linguagem natural. Os sistemas difusos propõem uma abordagem matemática para traduzir e processar o conhecimento subjectivo obtido de um especialista, ou seja, uma maneira de manipular conhecimento interpretável possuindo um determinado nível de incerteza. Esta incerteza é captada através dos conjuntos difusos que permitem a definição de um conjunto de regras que descrevem qualitativamente o comportamento dos sistemas.

Para uma melhor compreensão da forma como um modelo difuso descreve o comportamento de um dado sistema utilizando uma linguagem natural, interpretável, apresenta-se a descrição difusa de um exemplo muito simples já referido no capítulo 2. O exemplo consiste na modelização de um sistema de aquecimento através da relação estática entre a potência de aquecimento e a temperatura de saída.

A descrição deste modelo pode ser efectuada através do seguinte conjunto de regras do tipo SE-ENTÃO:

- *SE a potência de aquecimento é reduzida, ENTÃO a temperatura é baixa;*
- *SE a potência de aquecimento é média, ENTÃO a temperatura é moderada;*
- *SE a potência de aquecimento é elevada, ENTÃO a temperatura é alta;*

As regras qualitativas difusas utilizam variáveis linguísticas e termos, ou conjuntos, difusos na descrição do comportamento de um dado sistema. As variáveis linguísticas são as variáveis de entrada e saída que caracterizam um dado sistema, por exemplo, potência de aquecimento e temperatura. Os termos ou conjuntos difusos caracterizam uma dada variável linguística em vários subconjuntos como, por exemplo, os conjuntos difusos *elevada*, *média* e *reduzida*, para a variável linguística potência de aquecimento.

A utilização de variáveis linguísticas e conjuntos difusos na definição das regras permite incutir um determinado grau de incerteza idêntico ao utilizado na linguagem natural na descrição de um dado sistema. Este grau de incerteza é introduzido através da forma, e do posicionamento, dos conjuntos difusos no universo de variação das respectivas variáveis linguísticas.

As regras difusas permitem a utilização de conhecimento qualitativo de especialistas para a construção de modelos difusos com um elevado grau de interpretabilidade. O conjunto de regras difusas é designado por base de regras e é um dos blocos mais importantes da estrutura dos sistemas difusos.

Os sistemas difusos são constituídos, fundamentalmente por quatro blocos: fusificação, base de conhecimentos, inferência difusa e desfusão (figura 3.1):

Essencialmente, os sistemas difusos tornam possível usar categorias difusas como representação de palavras e ideias abstractas dos seres humanos para a descrição do procedimento de tomada de decisão.

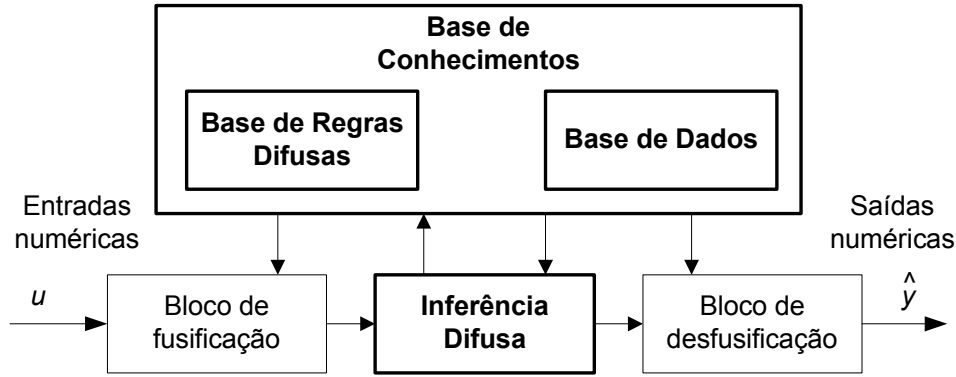


Figura 3.1: Estrutura de um sistema difuso.

### 3.2.1 Conjuntos Difusos

A teoria convencional de conjuntos baseia-se na premissa de que um elemento pertence, ou não, a um dado conjunto. A teoria dos conjuntos difusos toma uma posição difusa, no sentido de um dado elemento pertencer a mais do que um conjunto com um determinado grau de pertença. Tome-se a seguinte regra do exemplo atrás referido:

- *SE a potência de aquecimento é reduzida, ENTÃO a temperatura é baixa;*

Na teoria convencional de conjuntos, uma determinada temperatura é *baixa* se pertencer a um intervalo limitado como, por exemplo, o intervalo definido na equação 3.1:

$$baixa = [0; 20, 0[ \text{ } ^\circ C \quad (3.1)$$

Este tipo de conjuntos é definido por funções com variações abruptas. A representação convencional do conjunto *baixa* relativo à variável linguística temperatura não é muito correcta, uma vez que  $19,9 \text{ } ^\circ C$  pertence ao referido conjunto e  $20,0 \text{ } ^\circ C$  já não pertence ao referido conjunto, apesar de serem valores muito próximos.

Na teoria dos conjuntos difusos, as variáveis linguísticas variam numa determinada gama designada por universo de discurso  $U$ . Cada valor específico da variável linguística  $x$  pertence a um ou vários conjuntos difusos  $A(x)$ , com um determinado grau de pertença  $\mu_A(x)$  (equação 3.2):

$$\mu_A(x) \in [0; 1] \quad \text{em que} \quad x \in U \quad (3.2)$$

Os conjuntos difusos são assim definidos através de funções matemáticas com variações suaves que podem possuir várias formas tais como funções triangulares, trapezoidais, Gaussianas, do tipo sino, *singletons*, etc. [Wan97].

Estas funções de pertença tem um papel muito importante na representação correcta dos conjuntos difusos de uma dada variável, podendo estas serem funções lineares ou não lineares.

A função de pertença triangular, definida na equação 3.3, é um exemplo de uma função linear (definida por ramos lineares) bastante utilizada na representação de conjuntos difusos:

$$\mu_A(x; a; b; c) = \begin{cases} 0; & x \geq a \\ \left(1 - \frac{b-x}{b-a}\right); & a < x \leq b \\ \left(\frac{c-x}{c-b}\right); & b < x \leq c \\ 0; & x > c \end{cases} \quad (3.3)$$

em que  $a$ ,  $b$  e  $c$  são os vértices do triângulo, observando-o da esquerda para a direita.

A função pertença Gaussiana, definida na equação 3.4, é um exemplo de uma função não linear bastante utilizada na representação de conjuntos difusos:

$$\mu_A(x; c; \sigma) = e^{-\frac{(x-c)^2}{2\sigma^2}} \quad (3.4)$$

em que  $c$  e  $\sigma$  são o centro e o desvio padrão da função de pertença, respectivamente. Esta função é muito utilizada, uma vez que permite representar um dado conjunto difuso com um determinado grau de incerteza representado por uma função densidade de probabilidade muito conhecida.

Normalmente, as funções de pertença são simétricas relativamente ao seu centro; contudo, a representação de sistemas complexos não lineares pode exigir a definição de funções de pertença não simétricas para melhor representar as referidas não linearidades.

Além da forma, também o número de conjuntos difusos é um factor importante a definir. A definição do número de conjuntos difusos depende da resolução com que se deseja descrever a respectiva variável linguística (tipicamente, utilizam-se de 3 a 7 conjuntos difusos por cada variável linguística).

Na figura 3.2 podem-se observar os conjuntos difusos triangulares definidos para a variável linguística temperatura para o exemplo do sistema de aquecimento referido atrás. Cada valor numérico de temperatura irá pertencer a um, ou mais, conjuntos difusos com um determinado grau de pertença.

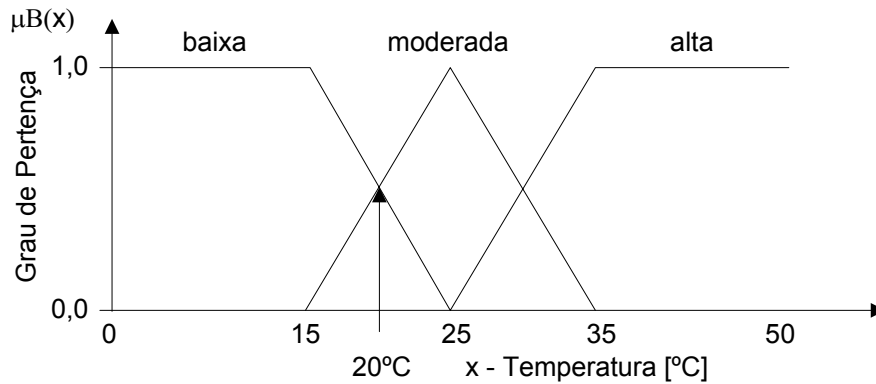


Figura 3.2: Definição dos conjuntos difusos da variável temperatura.

Sendo *baixa*, *moderada* e *alta* os conjuntos difusos da variável linguística temperatura, então para uma dada temperatura  $x$  os respectivos graus de pertença são



$\mu_{baixa}(x)$ ,  $\mu_{moderada}(x)$  e  $\mu_{alta}(x)$ . Segundo os conjuntos difusos triangulares definidos na figura 3.2, uma temperatura de 20.0°C possui os seguintes graus de pertinência  $\mu_{baixa}(20, 0) = 0,5$ ,  $\mu_{moderada}(20, 0) = 0,5$  e  $\mu_{alta}(20, 0) = 0,0$ .

### 3.2.2 Fusificação

A primeira etapa a realizar num sistema difuso é a fusificação das variáveis de entrada. Esta operação é responsável pela conversão das entradas numéricas em valores difusos, ou seja, cada entrada será descrita de uma forma difusa através dos vários conjuntos difusos. Com a definição das variáveis de entrada e os respectivos conjuntos difusos, cada valor numérico será transformado numa grandeza difusa caracterizada pelos seus graus de pertinência relativamente aos seus conjuntos difusos. Deste modo, é possível a conversão das entradas numéricas em linguagem natural com um determinado nível de incerteza que permitirá a construção de um conjunto de regras que irão descrever o comportamento do sistema.

### 3.2.3 Base de Conhecimentos

A base de conhecimentos é o “cérebro” dos sistemas difusos sendo constituída por uma base de dados e uma base de regras.

A base de dados contém informações acerca das gamas de variação das variáveis de entrada e saída, dos conjuntos difusos e das respectivas funções de pertinência.

A base de regras é constituída por um conjunto de regras do tipo SE-ENTÃO, constituídas por duas partes: as proposições dos antecedentes (premissas) e as proposições dos consequentes (consequências), como definido na seguinte regra genérica:

$$- SE <proposição de antecedentes>, ENTÃO <proposição de consequentes>;$$

As proposições dos antecedentes representam as relações possíveis entre os conjuntos difusos e as várias entradas. As proposições dos consequentes, representam as contribuições das relações das saídas de cada regra relativas às proposições dos respectivos antecedentes.

Dependendo da forma como as regras SE-ENTÃO são definidas, podem-se distinguir dois tipos de sistemas difusos: sistemas difusos do tipo Mandani (linguísticos) e sistemas difusos do tipo Takagi-Sugeno.

### Sistemas Difusos do Tipo Mandani

Os sistemas difusos do tipo Mandani [Man76], para o caso de um sistema MISO, possuem regras em que os antecedentes e os consequentes são proposições difusas, como na regra genérica definida na equação 3.5:

$$R_i : - SE \ x_1 \text{ é } A_{i1} \text{ e } \dots \text{ e } x_p \text{ é } A_{ip}; \ ENTÃO \ y_i \text{ é } B_i; \ i = 1..k \quad (3.5)$$

em que  $A_{i1}, \dots, A_{ip}$  e  $B_i$  são os termos linguísticos dos antecedentes e dos consequentes, respectivamente, para a regra  $i$  referente às entradas  $x_1, \dots, x_p$  e saída  $y_i$  em que  $k$  é o

número de regras que descrevem o sistema difuso. Neste caso, o sistema designa-se por sistema difuso linguístico, pois tanto as entradas como as saídas são representadas por termos linguísticos, obtendo regras completamente interpretáveis.

Este tipo de sistema difuso faz uma descrição totalmente linguística do sistema a modelizar, sendo tipicamente utilizado quando existem disponíveis conhecimentos experimentais provenientes de especialistas.

### Sistemas Difusos do Tipo Takagi-Sugeno

Os sistemas difusos do tipo Takagi-Sugeno [TS85], para o caso de um sistemas MISO, possuem regras em que os antecedentes são proposições difusas e os consequentes são funções lineares referentes às entradas, como expresso na regra genérica definida na equação 3.6:

$$R_i : - SE \ x_1 \text{ é } A_{i1} \text{ e } \dots \text{ e } x_p \text{ é } A_{ip}, \text{ ENTÃO } y_i = f_i(\mathbf{x}); \ i = 1..k \quad (3.6)$$

com

$$f_i(\mathbf{x}) = c_i, \quad i = 1..k \quad (3.7)$$

ou

$$f_i(\mathbf{x}) = \mathbf{a}_i^T \mathbf{x} + b_i, \quad \mathbf{x} = [x_1 \ x_2 \ \dots \ x_p]^T \quad i = 1..k \quad (3.8)$$

em que  $c_i$  é o vector das constantes dos consequentes,  $\mathbf{a}_i^T$  é o vector dos parâmetros dos consequentes relativamente às entradas do sistema,  $b_i$  é um vector que contém as componentes de ajuste estático e  $k$  é o número de regras que descrevem o sistema difuso.

Se a função  $f_i(\mathbf{x})$  for definida através da equação 3.7, então os consequentes são funções lineares de ordem zero (constantes). Se  $f_i(\mathbf{x})$  for definida através da equação 3.8, então os consequentes são funções lineares de ordem um dependentes das entradas.

Os sistemas difusos do tipo Takagi-Sugeno permitem a representação de um sistema global linear ou não linear através da contribuição linguística das entradas, reflectidas em funções lineares locais nos consequentes. Este tipo de sistemas difusos é utilizado quando os conhecimentos disponíveis dos especialistas são reduzidos, existindo disponíveis um conjunto representativo de dados de entrada e saída.

### 3.2.4 Inferência Difusa

Nos sistemas difusos a proposição dos antecedentes é representada pela conjugação dos vários graus de pertença de cada entrada num único grau de pertença global referente a cada regra  $\beta_i(\mathbf{x})$ . A proposição dos consequentes é também a conjugação dos vários graus de pertença de cada regra gerando um único valor final na saída do sistema.

As operações de conjugação dos vários conjuntos difusos dos antecedentes e dos consequentes são realizadas utilizando, tipicamente, os operadores de Norma-T e CoNorma-T, como os operadores lógicos “e” (conjunção) e “ou” (disjunção).

Na teoria difusa, vários operadores lógicos foram introduzidos para a realização destas operações como, por exemplo, os operadores de Zadeh e os operadores Probabilísticos [DHR93], definidos, na tabela 3.1.

	$A \text{ e } B$	$A \text{ ou } B$
Zadeh	$\min(\mu_{A_i}, \mu_{B_i})$	$\max(\mu_{A_i}, \mu_{B_i})$
Probabilístico	$\mu_{A_i} \mu_{B_i}$	$\mu_{A_i} + \mu_{B_i} - \mu_{A_i} \mu_{B_i}$

Tabela 3.1: Operadores lógicos difusos.

Para cada regra  $i$ , o cálculo do grau de pertença global proveniente dos antecedentes é efectuado utilizando um operador conjunção. O operador conjunção Zadeh calcula o grau de pertença global utilizando o operador mínimo como definido na equação 3.9:

$$\beta_i(\mathbf{x}) = \min(\mu_{A_{i1}}(x_1), \mu_{A_{i2}}(x_2), \dots, \mu_{A_{ip}}(x_p)) \quad (3.9)$$

O operador Probabilístico calcula o grau de pertença global utilizando o produto algébrico como definido na equação 3.10:

$$\beta_i(\mathbf{x}) = \mu_{A_{i1}}(x_1) \mu_{A_{i2}}(x_2) \dots \mu_{A_{ip}}(x_p) \quad (3.10)$$

### 3.2.5 Desfusificação

A última etapa de um sistema difuso é a desfusificação. Esta etapa é responsável pela conversão da saída difusa numa saída numérica. A desfusificação faz a conjugação dos vários graus de pertença globais de cada regra obtendo um valor difuso para a saída que, por sua vez, é transformado num valor numérico.

Nos sistemas difusos do tipo Takagi-Sugeno, a estimação da saída numérica  $\hat{y}$ , equação 3.11, é calculada através da média pesada da contribuição dos graus de pertença dos antecedentes com as funções lineares das várias regras:

$$\hat{y} = \frac{\sum_{i=1}^k \beta_i(\mathbf{x}) y_i}{\sum_{i=1}^k \beta_i(\mathbf{x})} = \frac{\sum_{i=1}^k \beta_i(\mathbf{x}) f_i(\mathbf{x})}{\sum_{i=1}^k \beta_i(\mathbf{x})} \quad i = 1..k \quad (3.11)$$

Tipicamente, os sistemas difusos do tipo Mandani estimam a saída numérica  $\hat{y}$ , combinando os vários conjuntos difusos de saída truncados pelos graus de pertença dos antecedentes, ou seja, unindo as várias sub-áreas dos vários conjuntos difusos sobrepostos das saídas e calculando, por exemplo, o seu centro de área. É de referir que existem muitos outros métodos de desfusificação como, por exemplo, o centro das somas, primeiro máximo e média dos máximos [DHR93].

Com base no exemplo do sistema de aquecimento, utilizando as regras definidas e os conjuntos difusos das variáveis de entrada e saída apresentados na figura 3.3, para

o caso de uma potência de aquecimento de entrada de 38%<sup>1</sup>, o valor numérico da saída calculado pelo método de desfusãoção centro de área (COA)[Vie97] é 14,33 °C.

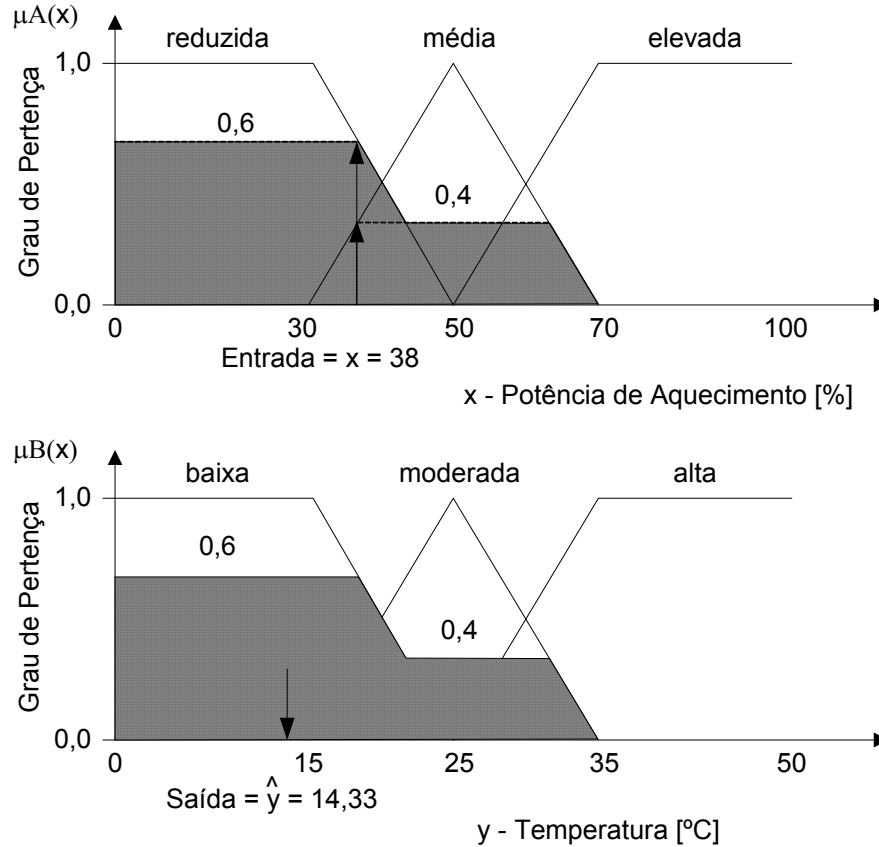


Figura 3.3: Aplicação do método de desfusãoção do centro de área a sistemas tipo Mandani, para o exemplo do sistema de aquecimento.

### 3.2.6 Comentários

De uma forma geral os modelos difusos permitem a descrição do comportamento de qualquer sistema através de um conjunto de regras escritas numa linguagem natural. Esta característica permite a inclusão de conhecimentos experimentais de especialistas de uma forma simples e intuitiva.

As vantagens mais significativas destes sistemas são:

- a capacidade de representar incertezas inerentes ao conhecimento humano utilizando variáveis linguísticas;
- a interacção simples do especialista da área com o projectista do sistema;

<sup>1</sup>As unidades de potência de aquecimento estão expressas em % e variam de 0 a 100%, correspondendo a potências de aquecimento mínima e máxima, respectivamente.

- a fácil interpretação dos resultados, devido à representação quase natural do conhecimento através das regras;
- a fácil extensão da base de conhecimento pela adição de novas regras, sem perda do conhecimento anterior;
- a criação de modelos, sem recorrer aos modelos fenomenológicos, utilizando apenas conhecimento empírico;
- a criação de modelos robustos relativamente a perturbações.

No entanto, estes sistemas são difíceis de ajustar sem recorrer a ajuda de algoritmos de aprendizagem numéricos. As desvantagens mais significativas destes sistemas são:

- incapazes de generalizar (ou seja, só responde ao que está descrito na base de regras);
- baixa robustez em relação a mudanças do sistema, uma vez que tais mudanças exigem alterações na base de regras;
- dependência da existência de um especialista para determinar as regras difusas;
- sensibilidade à subjectividade do especialista;
- dificuldade de optimização e ajuste.

### 3.3 Redes Neurais Artificiais

As redes neuronais tentam modelizar directamente as funções biológicas do cérebro humano. Isto leva à idealização dos neurónios como unidades discretas de processamento distribuído. As conexões locais ou globais de vários neurónios dentro de uma rede dão a estas estruturas a capacidade de assimilar, aprender, prever e/ou tomar decisões.

O primeiro modelo de rede neuronal foi descrito na década de 40 do século XX por W. S. McCulloch e W. Pitts e o primeiro algoritmo de treino foi desenvolvido, cerca de quinze anos mais tarde, por Hoff e B. Widrow [WH60].

A principal, e mais importante, característica das redes neuronais é a capacidade que estas estruturas têm de se ajustar ao sistema em causa através de um conjunto de exemplos (pares de dados de entradas e saídas). As redes neuronais modificam a sua estrutura interna, os pesos das conexões entre os seus neurónios artificiais, com o objectivo de mapear as relações existentes entre as entradas e saídas de um dado sistema.

As redes neuronais artificiais são estruturas constituídas por vários neurónios, definidos por funções matemáticas não lineares pesadas, colocadas em camadas, interligadas entre si. Tipicamente, as redes neuronais artificiais possuem vários neurónios ligados em série e em paralelo em que o fluxo da informação segue da entrada para a saída. Este tipo de rede neuronal artificial é chamado *Feedforward Multilayer Neural Network* (FMNN), ou seja, rede neuronal directa multi-camada.

### 3.3.1 Rede Neuronal Multi-Camada

As redes neuronais artificiais multi-camada são estruturas com várias camadas interligadas através de pesos, camadas essas que possuem elementos simples de processamento chamados neurónios. A informação relevante relativa ao mapeamento da entrada e saída fica armazenada nesses pesos.

Para efeitos ilustrativos, considere-se a FMNN, apresentada na figura 3.4, com  $p$  entradas,  $n$  saídas e com uma camada escondida.

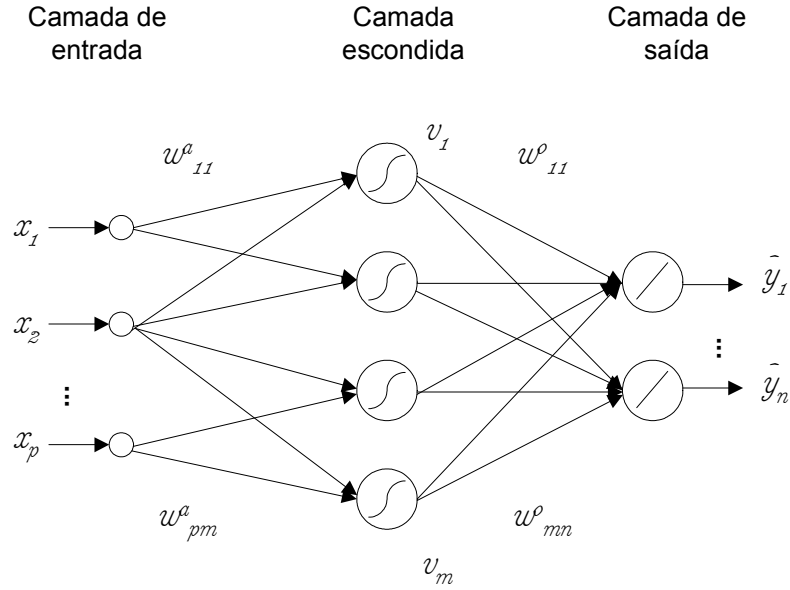


Figura 3.4: Rede neuronal artificial com uma camada escondida ( $p$  entradas,  $n$  saídas e  $m$  neurónios na camada escondida).

A camada de entrada não efectua qualquer operação, distribuindo apenas as entradas  $x_i$  para os pesos  $w_{ij}^a$  da camada escondida. Em cada neurónio da camada escondida a soma pesada das entradas é calculada, como definido na equação 3.12:

$$z_j = \sum_{i=1}^p w_{ij}^a x_i = (\mathbf{w}_j^a)^T \mathbf{x}, \quad j = 1..m \quad (3.12)$$

em que  $a$  e  $o$  referem-se aos pesos da camada escondida e de saída, respectivamente, e  $i$  e  $j$  referem-se aos pesos relativos às entradas/saídas e ao respectivo neurónio da camada escondida.

O resultado anterior é passado por uma função de activação não linear como a função tangente hiperbólica definida na equação 3.13:

$$v_j = \frac{1 - e^{(-2z_j)}}{1 + e^{(-2z_j)}}, \quad j = 1..m \quad (3.13)$$

Finalmente, os neurónios da camada de saída têm uma função de activação linear, calculando a soma pesada das suas entradas, como definido na equação 3.14, obtendo-se

a saída do sistema:

$$\hat{y}_l = \sum_{i=1}^a w_{jl}^a v_l = (\mathbf{w}_j^a)^T \mathbf{v}, \quad l = 1..n \quad (3.14)$$

O treino destas estruturas consiste na adaptação dos pesos da rede de forma a minimizar o erro entre a saída da rede e a saída do sistema. Este procedimento é realizado em duas fases. Primeiro, calcula-se a saída da rede, calculando as saídas dos neurónios camada após camada até atingir a saída. Em segundo, faz-se a adaptação dos pesos, calculando as derivadas parciais entre a saída desejada e a saída da rede (erro), propagando-as para as camadas internas da rede de forma a ajustar os respectivos pesos. Esta propagação é realizada da saída para a entrada. As derivadas parciais erro são propagadas para a camada de saída, de seguida para a camada anterior até atingir a (primeira) camada escondida de modo a ajustar todos os pesos da rede. Esta técnica de propagação do erro é conhecida por retropropagação do erro.

A utilização de FMNN com uma camada escondida é suficiente para a modelização da maioria dos sistemas [BV03]. Um maior número de camadas pode fornecer um menor erro de modelização; em contrapartida, aumenta a complexidade da rede e o seu número de parâmetros. A escolha do número de neurónios da camada escondida é essencial para a criação de modelos com baixo erro de modelização. Poucos neurónios implicam um elevado erro de modelização, muitos neurónios resultam numa rede com muitos pesos, potenciando uma fraca capacidade de generalização. Um compromisso entre estas duas situações deve ser levado em consideração na construção de uma FMNN.

### 3.3.2 Método de Retropropagação do Erro

Para maior simplicidade na explicação do método da retropropagação do erro, considera-se uma FMNN com  $p$  entradas e uma saída e um conjunto de  $N$  pares de dados entrada e saída  $\{\mathbf{X}, \mathbf{y}\}$ , definidos na equação 3.15:

$$\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N]; \quad \mathbf{y} = [y_1, \dots, y_N]^T \quad (3.15)$$

em que  $\mathbf{X} \in \mathbb{R}^{N \times p}$  é um vector linha que contem as  $p$  entradas  $\mathbf{x}_k$  e  $\mathbf{y} \in \mathbb{R}^N$  é um vector coluna que contem as saídas desejadas  $y_k$  com  $k = 1..N$ .

Chamando  $e$  à diferença entre a saída desejada  $y$  e a saída da rede  $\hat{y}$ , o ajuste dos pesos da rede é conseguido minimizando a função de custo definida na equação 3.16:

$$J = \frac{1}{2} \sum_{k=1}^N e_k^2 \quad \text{com} \quad e_k = y_k - \hat{y}_k \quad (3.16)$$

De notar que a saída da rede é uma função não linear relativamente aos seus pesos  $\mathbf{w}$ . Assim, o treino de uma FMNN é um problema de optimização não linear para o qual vários métodos baseados na retropropagação do erro podem ser aplicados:

- Método do Gradiente (gradiente de primeira ordem);

- Método de Newton (gradiente de segunda ordem);
- Método de Levenberg-Marquard (gradiente de segunda ordem).

Os métodos que utilizam o gradiente de primeira ordem baseiam-se na regra de actualização dos pesos definida na equação 3.17:

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \lambda(n) \nabla J(\mathbf{w}(n)) \quad (3.17)$$

em que  $\mathbf{w}(n)$  é o vector dos pesos na iteração  $n$ ,  $\lambda(n)$  é o factor de aprendizagem e  $\nabla J(\mathbf{w}(n))$  é o Jacobiano da FMNN definido na equação 3.18:

$$\nabla J(\mathbf{w}(n)) = \left[ \frac{\partial J(\mathbf{w})}{\partial w_1}, \frac{\partial J(\mathbf{w})}{\partial w_2}, \dots, \frac{\partial J(\mathbf{w})}{\partial w_M} \right]^T \quad (3.18)$$

O problema da optimização não linear é resolvido usando o primeiro termo da expansão de  $J(\mathbf{w})$  em série de Taylor:

$$J(\mathbf{w}) \approx J(\mathbf{w}_0) + \nabla J(\mathbf{w}_0)^T (\mathbf{w} - \mathbf{w}_0) + \frac{1}{2} (\mathbf{w} - \mathbf{w}_0) \mathbf{H}(\mathbf{w}_0) (\mathbf{w} - \mathbf{w}_0) \quad (3.19)$$

Os algoritmos de aprendizagem que utilizam o gradiente de segunda ordem são mais eficientes do que os que utilizam o gradiente de primeira ordem, uma vez que esta abordagem aumenta a velocidade de convergência do algoritmo de aprendizagem.

Os métodos que utilizam o gradiente de segunda ordem, além do primeiro termo, utilizam também o segundo termo da expansão de  $J(\mathbf{w})$  em série de Taylor, equação 3.19, em que  $\mathbf{H}(\mathbf{w}_0)$  é a matriz Hessiana que contém as derivadas de segunda ordem de  $J$  relativamente aos vários pesos.

### 3.3.3 Comentários

As redes neuronais artificiais permitem aproximar qualquer sistema através de um conjunto de neurónios dispostos em camadas interligadas por pesos, cuja identificação é efectuada à custa de um conjunto de dados de entrada e saída do sistema.

A interpretabilidade destas estruturas é, praticamente nula, uma vez que a informação sobre o sistema está armazenada de forma distribuída. Este facto faz com que este tipo de modelos seja classificado como modelos do tipo caixa negra.

As vantagens mais importantes das redes neuronais são:

- a capacidade de aprendizagem;
- a capacidade de generalização: responde correctamente a novos casos desde que esses sejam semelhantes aos casos aprendidos;
- a robustez em relação a perturbações.

As desvantagens mais importantes são:

- a ausência de interpretabilidade dos modelos criados;
- a dificuldade em definir a respectiva estrutura interna (número de camadas e número de neurónios).



## 3.4 Sistemas Neuro-Difusos

Desde que os sistemas difusos se tornaram uma aplicação comum no meio industrial, os projectistas perceberam que o desenvolvimento de um sistema difuso que apresente um bom desempenho não era uma tarefa fácil. O problema de encontrar funções de pertença e regras apropriadas era, frequentemente, um processo cansativo de tentativa e erro. Assim, surgiu a ideia de aplicar algoritmos de aprendizagem aos sistemas difusos. As técnicas de aprendizagem utilizadas nas redes neuronais surgem como uma solução adequada para automatizar, ou dar suporte, ao desenvolvimento de sistemas difusos ajustáveis de forma automática.

As redes neuronais e os sistemas difusos podem ser integrados ou combinados para unirem as suas vantagens e sanarem as suas deficiências individuais. As redes neuronais introduzem as suas características computacionais de aprendizagem nos sistemas difusos e recebem destes o poder de representação da informação de forma interpretável. Deste modo as desvantagens dos sistemas difusos são compensadas pelas capacidades das redes neuronais e as vantagens dos sistemas difusos contornam as limitações das redes neuronais. As técnicas são complementares, o que justifica a sua combinação na solução de problemas aproveitando ao máximo as potencialidades de cada uma delas evitando as respectivas limitações.

Na definição de Nauck e Kruse [NKK97]: “Um sistema neuro-difuso é um sistema difuso que usa um algoritmo de aprendizagem com base nas derivadas ou inspirado pela teoria de redes neuronais para determinar os seus parâmetros (parâmetros dos conjuntos difusos e/ou regras difusas) através da utilização de dados de treino (dados de entrada e saída)”.

Os sistemas resultantes da fusão de sistemas difusos e redes neuronais tem como vantagens a aprendizagem através de dados de treino (característica das redes neuronais) e a fácil interpretação funcional do sistema (característica dos sistemas difusos).

### 3.4.1 Construção de um Sistema Neuro-Difuso

Ambos os tipos de informação, conhecimentos empíricos e experimentais e dados de entrada e saída, podem ser utilizados para a construção de sistemas neuro-difusos. Normalmente, o conhecimento *a priori* tem uma natureza mais aproximada (conhecimentos qualitativos, empíricos e interpretáveis), enquanto que os dados de entrada e saída possuem informação mais precisa (conhecimento quantitativo, determinístico e não interpretável).

A inicialização da estrutura dos sistemas neuro-difusos pode ser efectuada de duas formas diferentes dependendo do tipo de informação inicial existente sobre o sistema. A primeira forma consiste em inicializar a estrutura com base num conjunto de conhecimentos especializados e experimentais que permitem a definição de uma colecção de regras do tipo SE-ENTÃO, assim como, a definição das variáveis linguísticas significativas e respectivos conjuntos difusos. A segunda consiste na aplicação de métodos de partição do espaço de entrada ou agrupamentos de dados que se baseiam em dados de entrada e saída, permitindo a definição de um conjunto de regras do tipo SE-ENTÃO, assim como, a definição da forma e posicionamento dos respectivos conjuntos difusos.

Ao contrário das redes neuronais artificiais, os sistemas neuro-difusos permitem a geração um conjunto de regras interpretáveis ou com algum nível de interpretabilidade que descrevem o comportamento de um dado sistema real.

Com base na combinação dos vários tipos de conhecimentos utilizados, é possível definir e identificar os dois tipos de sistemas neuro-difusos: sistemas neuro-difusos do tipo Mandani e do tipo Takagi-Sugeno.

### Sistemas Neuro-Difusos do tipo Mandani

Os sistemas neuro-difusos do tipo Mandani são baseados em sistemas difusos do tipo Mandani. Na figura 3.5 apresenta-se, como exemplo ilustrativo, a estrutura de um sistema neuro-difuso do tipo Mandani com  $p$  entradas  $x$  e uma saída  $\hat{y}$ .

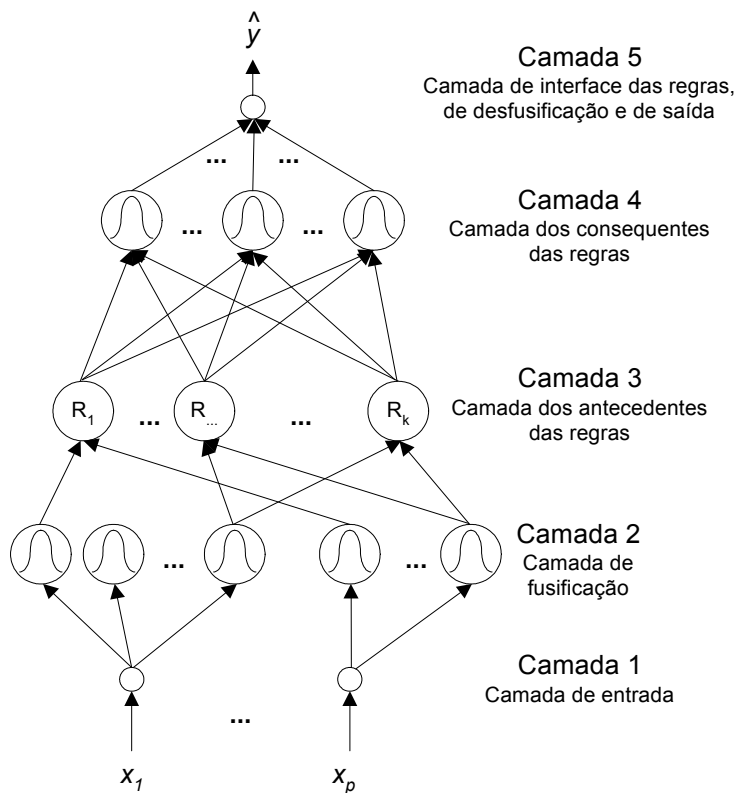


Figura 3.5: Sistema neuro-difuso do tipo Mandani.

Este sistema é constituído por cinco camadas interligadas entre si em que cada camada realiza uma dada função específica, a saber:

**Camada 1** (camada de entrada):

- Nesta camada não são efectuados quaisquer cálculos. Cada nó desta camada, que corresponde a uma variável de entrada, transmite os valores da entrada para a camada seguinte.

**Camada 2** (camada de fusificação):

- Cada nó desta camada corresponde a um dado conjunto difuso (por exemplo: *baixa*, *moderada*, *alta*) relativo a uma dada variável de entrada da camada 1, ou seja, as saídas desta camada representam os valores de pertença que especificam os graus com que os valores de entrada pertencem aos respectivos conjuntos difusos.

No caso da utilização da função de pertença Gaussiana, definida na equação 3.20:

$$\mu_{A_{il}}(x_l; c_{il}; \sigma_{il}) = e^{-\frac{(x_l - c_{il})^2}{2\sigma_{il}^2}} \quad (3.20)$$

em que  $c_{il}$  e  $\sigma_{il}$  são o centro e o desvio padrão da referida função, os graus de pertença são dados por  $\mu_{A_{il}}$  relativos a um dado conjunto difuso  $A_{il}(\cdot)$ , a uma regra  $i$  e a uma entrada  $x_l$ ,  $l = 1..p$ .

A inicialização do número e o tipo das funções de pertença a atribuir a cada variável pode ser definido utilizando conhecimentos *a priori*, utilizando algoritmos de agrupamento de dados ou métodos de partição do espaço de entrada (descritos na secção seguinte). A forma final das funções de pertença será ajustada durante o processo de aprendizagem da rede (descrito na secção 3.5).

**Camada 3** (camada dos antecedentes das regras):

- Cada nó desta camada representa a proposição dos antecedentes ou premissas de cada regra. Tipicamente, cada nó usa um operador do tipo Norma-T (associado à intercepção de conjuntos) que calcula o grau de pertença global dos antecedentes de cada regra. Este grau de pertença global depende dos graus de pertença dos conjuntos difusos dos antecedentes de cada entrada.

Se utilizar como operador Norma-T o produto algébrico, o grau de pertença global de cada regra  $i$  é calculado como definido na equação 3.21:

$$\beta_i = \mu_{A_{i1}}(x_1; c_{i1}; \sigma_{i1}) \mu_{A_{i2}}(x_2; c_{i2}; \sigma_{i2}) \dots \mu_{A_{ip}}(x_p; c_{ip}; \sigma_{ip}), \quad i = 1, 2, \dots, k \quad (3.21)$$

**Camada 4** (camada dos consequentes das regras):

- Nesta camada cada nó efectua duas tarefas. Combina os graus de pertença dos antecedentes das regras vindos da camada anterior e determina os graus de pertença relativos aos conjuntos difusos dos consequentes. O número de nós desta camada será igual ao número de regras.

A inicialização do número e do tipo das funções de pertença dos consequentes das regras pode ser definido utilizando conhecimentos experimentais ou utilizando algoritmos de agrupamento de dados (descrito na secção seguinte). A forma final das funções de pertença é ajustada durante o processo de aprendizagem do sistema neuro-difuso (descrito na secção 3.5).

**Camada 5** (camada de interface das regras e de desfusão para uma saída):

- Cada nó desta camada corresponde a uma saída do sistema e em cada nó é efectuada a combinação de todos os consequentes das regras usando um operador do tipo CoNorma-T (associado à união de conjuntos) e, finalmente, é calculado o valor numérico de saída.

### Sistemas Neuro-Difusos do tipo Takagi-Sugeno

Estes sistemas são baseados em sistemas difusos do tipo Takagi-Sugeno. Na figura 3.6 apresenta-se, como exemplo ilustrativo, a estrutura de um sistema neuro-difuso do tipo

Takagi-Sugeno de ordem um com  $p$  entradas  $x$  e uma saída  $\hat{y}$ .

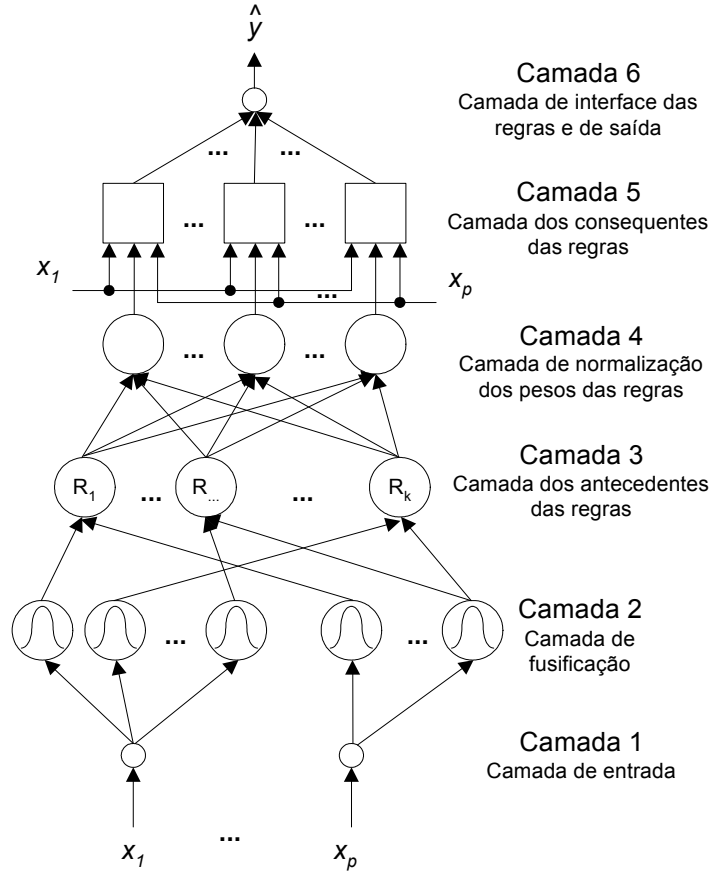


Figura 3.6: Sistema neuro-difuso do tipo Takagi-Sugeno.

Este sistema é constituído por seis camadas interligadas entre si em que cada camada realiza uma dada função específica, a saber:

As funções realizadas nas camadas 1, 2 e 3 dos sistemas neuro-difusos do tipo Takagi-Sugeno são idênticas às funções realizadas nas camadas 1, 2 e 3 dos sistemas do tipo Mandani.

**Camada 4** (camada de normalização do peso global das regras):

- Cada nó desta camada calcula a relação do peso de cada regra  $i$  relativamente à soma dos pesos de todas as regras, sendo este definido pela equação 3.22:

$$\bar{\beta}_i = \frac{\beta_i}{\beta_1 + \beta_2 + \dots + \beta_k}, \quad i = 1, 2, \dots, k \quad (3.22)$$

**Camada 5** (camada dos consequentes das regras):

- Todos os nós desta camada realizam a função definida na equação 3.23:

$$\bar{\beta}_i f_i = \bar{\beta}_i (p_i x_1 + q_i x_2 + \dots + t_i x_p + v_i), \quad i = 1, 2, \dots, k \quad (3.23)$$

onde  $\beta_i$  é a saída da camada 4,  $\{p_i, q_i, \dots, t_i, v_i\}$  são os parâmetros lineares dos consequentes e  $x_1, x_2, \dots, x_p$  são as entradas do sistema.

Tipicamente, o método utilizado para o cálculo dos parâmetros  $\{p_i, q_i, \dots, t_i, v_i\}$  é o método de mínimos quadrados.

**Camada 6** (camada de interface das regras):

- Cada nó nesta camada corresponde a uma saída do sistema. Cada nó faz a combinação de todos os consequentes das regras (funções locais lineares pesadas) somando todos os sinais que entram neste nó, como definido na equação 3.24 calculando desta forma o valor numérico de saída:

$$\hat{y} = \sum_{i=1}^k \overline{\beta_i} f_i = \sum_{i=1}^k \frac{\beta_i f_i}{\sum_{i=1}^k \beta_i}, \quad i = 1, 2, \dots, k \quad (3.24)$$

### 3.4.2 Estrutura e Parâmetros de um Sistema Neuro-Difuso

A definição da estrutura e respectivos parâmetros internos é um passo importante na construção de um sistema neuro-difuso. A estrutura destes sistemas depende do número de variáveis de entrada e saída, do número e tipo das funções de pertinência e do número de regras. A estrutura de um sistema neuro-difuso condiciona a flexibilidade com que este aproxima uma dada função. Sistemas com um reduzido número de parâmetros podem originar uma má aproximação; por outro lado, sistemas com um elevado número de parâmetros podem originar uma boa aproximação, mas aumentam a complexidade matemática da estrutura podendo reduzir a sua capacidade de generalização.

O processo de definição inicial da estrutura dos sistemas neuro-difusos envolve, essencialmente, as seguintes escolhas: selecção das variáveis de entrada, número/tipo das funções de pertinência e número/definição das regras difusas.

#### Seleccção das variáveis de entrada

A selecção das variáveis de entrada envolve a escolha das amostras das variáveis de entrada e saída atrasadas (regressores), que permitem a introdução da dinâmica (caso de um sistema dinâmico) do sistema no modelo. Os conhecimentos fenomenológicos e experimentais do sistema são, tipicamente, a base de informação mais utilizada para a escolha do conjunto inicial dos regressores.

Os métodos de escolha dos regressores referidos na literatura dividem-se em dois grupos principais: os métodos baseados no modelo e os *model-free methods*, ou seja, métodos livres do modelo [SB04]. Os métodos baseados no modelo constroem vários modelos com as várias combinações alternativas de regressores de entrada. Os modelos com menor erro de modelização indicam quais os regressores mais significativos na modelização do respectivo sistema. A escolha dos regressores nos métodos livres do modelo é baseada em métodos estatísticos, propriedades das funções, etc.

Dos métodos baseados em modelo destacam-se o método *Akaike's Information Criterion* (AIC), usado frequentemente nos sistemas lineares, e o método proposto por Jang [Jan96] para a identificação de modelos baseados nos sistemas neuro-difusos ANFIS [Jan93]. Este último inicializa e gera um conjunto de modelos que cobrem todos os possíveis regressores candidatos. Após uma primeira iteração de treino, só os modelos que atingem menores valores do RMSEs são reutilizados numa segunda iteração, uma

vez que os respectivos regressores possuem um maior potencial para a diminuição dos RMSEs em iterações futuras.

Dos métodos livres do modelo destacam-se o método apresentado em [HA93], que explora as propriedades da continuidade das funções não lineares e o método apresentado em [SB04], que determina os regressores mais significativos utilizando conceitos de geometria e aplicando as propriedades básicas de funções.

### Número/forma das funções de pertença e número/definição de regras

O número das funções de pertença e a definição do número de regras estão mutuamente relacionados (a introdução de mais funções de pertença implica a definição de um maior número de regras) e determinam o nível de detalhe de descrição do sistema (granularidade). A definição do número e da forma das funções de pertença e a definição e o número de regras pode ser baseada em conhecimentos experimentais. Na ausência deste tipo de conhecimentos, a sua definição pode ser baseada em dados de entrada e saída, utilizando métodos como: a partição do espaço de entrada em grelha ou em árvore ou utilizando o método de agrupamento de dados [JSM97]. Existem ainda métodos de construção/simplificação do conjunto de regras que se baseiam no peso que cada regra tem no modelo final [AN00].

**Partição do espaço de entrada** Para a partição do espaço de entrada existem vários métodos possíveis dos quais se destacam os seguintes: partição em grelha uniforme e partição em zonas não uniformes.

A partição uniforme do espaço de entrada consiste, como se ilustra a figura 3.7, na divisão do domínio das variáveis de entrada em zonas uniformemente distribuídas com funções de pertença de tipo e forma pré-definidas.

Este método inicializa uniformemente o espaço de entrada permitindo a definição das regras, sendo o número de conjuntos difusos e a forma inicial destes definidos *a priori*. No caso da figura 3.7, o número de regras é igual a nove, em que três é o número de funções de pertença atribuído a cada entrada e dois é o número total de entradas. Contudo, para sistemas complexos, com o aumento do número de entradas, a aplicação deste método conduz a um crescimento exponencial do número de regras, bem como do número de conjuntos difusos, o que implica estruturas relativamente complexas.

A partição uniforme nem sempre corresponde à melhor partição do espaço das entradas, surgindo outros tipos de partição como, por exemplo, a partição não uniforme.

A partição não uniforme do espaço de entrada consiste, como mostra a figura 3.8, na divisão do domínio das variáveis de entrada em zonas não uniformemente distribuídas com funções de pertença de tipo e forma pré-definidas.

A partição do domínio das variáveis de entrada em zonas não uniformes é definida recorrendo a algoritmos de procura discretos como o algoritmo CART (*Classification and Regression Tree*) [BFOS84] cuja procura é realizada por regressão em árvore. Este algoritmo tem duas partes: uma inicial chamada crescimento em árvore e, posteriormente, a poda da árvore. O crescimento da árvore consiste, inicialmente na partição dos dados de treino em dois grupos que minimizem uma dada função de erro. A esses dois grupos aplica-se uma nova subdivisão e assim sucessivamente, diminuindo cada

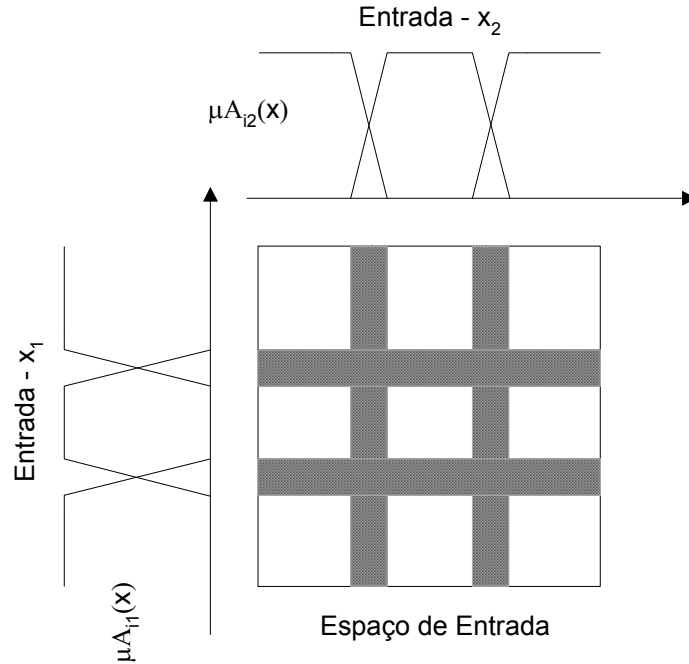


Figura 3.7: Partição uniforme do espaço de entrada (para o caso de um modelo com 2 entradas e funções de pertinência trapezoidais).

vez mais a referida função de erro até atingir um valor erro mínimo pré-definido. A poda da árvore consiste em retirar pequenos grupos gerados pelo método anterior que não tenham muita influência na partição final do espaço de entradas. A utilização de uma função de custo proporcional à complexidade da estrutura [JSM97] permite a simplificação da partição final.

Outro algoritmo em árvore, proposto em [NFBS00], vai fazendo partições, dividindo o espaço inicial ao meio, de seguida divide a partição que provoua um maior RMSE em duas partes, e assim sucessivamente, dividindo sempre o subconjunto ou partição que localmente apresente um maior RMSE, até que o erro atinja um determinado valor pré-definido. Este tipo de algoritmo pode conduzir a um número excessivo de regras, uma vez que a partição gerada pode não ser a mais adequada [NFBS00].

**Métodos de agrupamento de dados** Os métodos de agrupamento de dados permitem agrupar os pares de dados com características idênticas em vários subgrupos que são caracterizados por um centro e um raio de abrangência. Este métodos procuram agrupar os pares de dados num determinado grupo com características semelhantes minimizando uma determinada função de custo, normalmente relacionada com as distâncias dos dados aos respectivos centros (minimizar a distância dos dados aos respectivos centros).

Tipicamente, os algoritmos de agrupamento de dados originam uma partição dispersa no espaço dos dados de entrada. Os grupos gerados podem sobrepor-se ou não, dependendo das características dos dados e do número de grupos desejados, como ilus-

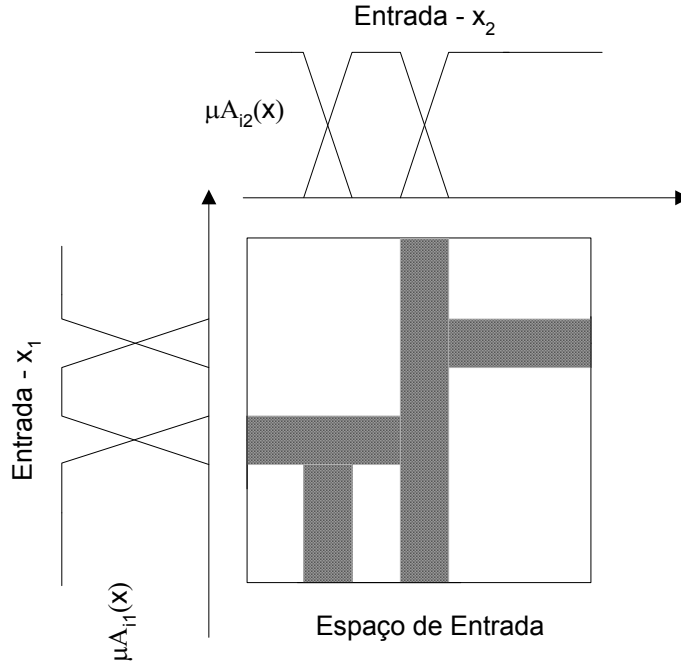


Figura 3.8: Partição não uniforme do espaço de entrada (para o caso de 2 entradas com funções de pertinência trapezoidais).

tra a figura 3.9.

No caso dos sistemas neuro-difusos, os métodos de agrupamento de dados podem servir para definir o número de regras, tal como o número e a forma das funções de pertinência utilizadas (por exemplo, o centro e variância no caso das funções de pertinência gaussianas muito utilizadas neste tipo de sistemas). A forma do agrupamento depende do tipo de distância utilizada na criação dos grupos, assim como da forma como os dados pertencem aos vários grupos.

Os métodos de agrupamento de dados subdividem-se em dois tipos: sem sobreposição e com sobreposição de dados. O agrupamento de dados que admite sobreposição utiliza funções de pertinência típicas dos sistemas difusos para a definição dos diferentes graus de pertinência dos dados aos respectivos grupos. Estes métodos de agrupamento de dados são designados como agrupamentos de dados difusos.

Dos métodos de agrupamento de dados sem sobreposição destacam-se: o agrupamento K-média (*K-Mean* - KM) e o agrupamento subtrativo.

O método de agrupamento de dados K-média (também conhecido por agrupamento C-média [KK82]) consiste na divisão dos dados em  $c$  grupos ( $G_i, i = 1..c$ ) e na procura dos centros dos grupos que minimizam uma função de custo. Esta função de custo é baseada numa medida da distância dos dados num dado grupo como, por exemplo, a distância euclidiana. Cada par de dados do referido espaço pertence a um único grupo de uma forma binária.

O método de agrupamento de dados subtrativo [Chi94] consiste no cálculo dos centros de uma forma sequencial baseada na medida da maior densidade de dados



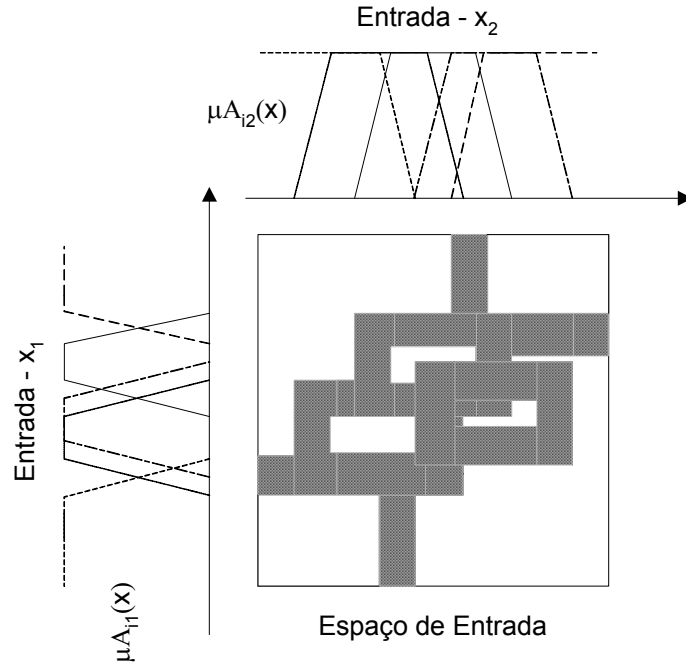


Figura 3.9: Partição dispersa do espaço de entrada (para o caso de 2 entradas com funções de pertinência trapezoidais).

de grupos esféricos com um determinado raio. Este processo é repetido até que um número suficiente de grupos seja gerado. Posteriormente, os centros destes grupos serão os centros das funções de pertinência a definir para construção do sistema neuro-difuso.

Dos métodos de agrupamento difusos que admitem sobreposição destacam-se: o agrupamento C-média difuso (*Fuzzy C-Mean* - FCM) e agrupamento difuso elipsoidal.

O método de agrupamento de dados C-média difuso [Bez73] é muito semelhante ao método de agrupamento de dados K-média, no entanto, neste caso cada par de dados pertence a um grupo com um determinado grau definido por uma função de pertinência difusa contido no intervalo  $[0; 1]$ . O número de grupos é fixo, calculando os centros e as variâncias que minimizam a função de custo baseada na soma das distâncias euclidianas dos vários pares de dados a um dado centro.

O método de agrupamento difuso elipsoidal [BV97] é um método de agrupamento de dados especialmente pensado para classificação de dados entrada e saída de um sistema dinâmico. Este método agrupa os dados em conjuntos com uma forma elipsoidal em vez da forma esférica típica da maioria dos métodos de agrupamento de dados. Este método define um conjunto de elipsóides que representam uma versão inicial do mapeamento entre as entradas e as saídas do sistema.

Outro método de agrupamento que utiliza os dados de entrada e saída é o agrupamento C-média difuso condicional [Ped96], semelhante ao método de agrupamento difuso de dados C-média, mas que toma em consideração os dados de saída para a definição dos grupos.

**Simplificação do número de regras** O conjunto de regras de um modelo neuro-difuso pode ser determinado com base nos seguintes métodos: redução ou adição. O método de redução consiste em retirar uma determinada regra da base de conhecimentos identificando novamente o modelo, se este continuar a obter baixos erros de modelização, esta regra poderá ser eliminada, caso contrário a regra é mantida. Este procedimento é executado para todas as regras pré-definidas [NK97] simplificando o modelo. O método de adição baseia-se na adição de regras, uma de cada vez, estudando o seu peso para o modelo global. Se o erro de modelização global diminuir com a integração desta regra, esta passa a fazer parte da base de conhecimentos, caso contrário é rejeitada. Este procedimento é executado até se atingir um determinado número de regras pré-definido ou até se obter um determinado erro de modelização pré-definido. Estes métodos fornecem uma ajuda muito importante para a simplificação e definição do número óptimo de regras.

Com o objectivo de manter, ou até aumentar, a interpretabilidade dos sistemas neuro-difusos, permitindo a simplificação do número de funções de pertença, Paiva [Pai99] e Babuska [Bab98] propõem técnicas de fusão de funções de pertença. Estas técnicas baseiam-se na transformação de dois ou mais conjuntos difusos que se encontram muito sobrepostos num único conjunto difuso. Outro aspecto importante na definição das regras neuro-difusas do tipo Takagi-Sugeno apresentado em [Bab98] é a imposição de funções lineares locais interpretáveis nos consequentes, uma vez que estas funções, nos sistemas neuro-difusos do tipo Takagi-Sugeno genéricos, são tipicamente não interpretáveis. Estas técnicas permitem reduzir o número de regras e de funções de pertença, contudo, podem diminuir a capacidade de aproximação final do respectivo modelo. Assim, deverá existir um compromisso entre a diminuição da complexidade da estrutura e a capacidade de aproximação desejada.

### 3.5 Identificação de Sistemas Neuro-Difusos

A identificação de sistemas neuro-difusos pode ser efectuada em linha e ou fora de linha. Este trabalho restringiu-se à identificação de sistemas neuro-difusos fora de linha utilizando conhecimentos experimentais e um conjunto de dados entrada e saída.

Existem vários algoritmos de identificação fora de linha de sistemas neuro-difusos. Se a respectiva estrutura for constituída por parâmetros exclusivamente não lineares, o algoritmo utilizado mais frequentemente é baseado na descida do gradiente tanto para o ajuste dos parâmetros dos antecedentes como para o ajuste dos parâmetros dos consequentes. Este algoritmo de identificação é, usualmente, utilizado no ajuste dos parâmetros internos de sistemas neuro-difusos do tipo Mandani.

Por outro lado, se a estrutura neuro-difusa contiver parâmetros lineares e não lineares, poder-se-ão utilizar algoritmos híbridos de aprendizagem normalmente mais eficientes. Estes algoritmos híbridos utilizam duas técnicas complementares para a determinação e ajuste dos referidos parâmetros. Se os parâmetros a ajustar forem não lineares, utilizam-se métodos baseados na descida do gradiente; se os parâmetros a ajustar forem lineares, utilizam-se métodos de identificação lineares tais como o método de mínimos quadrados. Este método de identificação híbrido é tipicamente utilizado na

identificação dos parâmetros internos de sistemas neuro-difusos do tipo Takagi-Sugeno.

### 3.5.1 Aprendizagem Baseada no Gradiente Descendente

Neste caso, qualquer que seja o tipo de sistema neuro-difuso, todos os parâmetros internos dos antecedentes e dos consequentes são identificados e ajustados através do método do gradiente de primeira ou segunda ordem. Os métodos baseados no gradiente de primeira ordem, conhecidos por métodos da retropropagação do erro, são os métodos mais vulgarmente utilizados devido à sua eficiência e baixa complexidade matemática. A aplicação de métodos baseados no gradiente de segunda ordem implica algoritmos de elevada complexidade matemática contudo permite maiores velocidades de convergência. Este procedimento de identificação é semelhante ao utilizado na identificação de redes neuronais (secção 3.3).

Para demonstrar a aplicação deste algoritmo, utilizou-se um sistema MISO neuro-difuso do tipo Takagi-Sugeno de ordem zero, com  $p$  entradas  $x_{1..p}$  e uma saída  $\hat{y}$ , como se vê na figura 3.10:

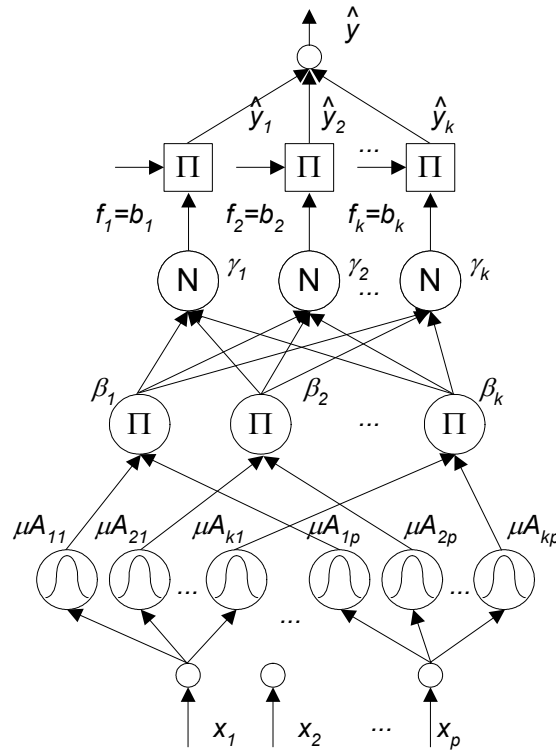


Figura 3.10: Sistema neuro-difuso do tipo Takagi-Sugeno de ordem zero com  $p$  entradas e uma saída.

Este sistema é representado por um conjunto de  $k$  regras difusas do tipo Takagi-Sugeno de ordem zero definidas na equação 3.25:

$$R_i : - SE \ x_1 \text{ é } A_{i1} \ x_2 \text{ é } A_{i2} \ e... \ x_p \text{ é } A_{ip}, \text{ ENTÃO } \hat{y}_i = v_i; \quad i = 1..k \quad (3.25)$$

em que as funções de pertença dos antecedentes são funções com transições suaves, como é o exemplo da função de pertença Gaussiana, definida através da equação 3.26:

$$\mu_{A_{il}}(x_l; c_{il}; \sigma_{il}) = e^{-\frac{(x_l - c_{il})^2}{2\sigma_{il}^2}}, \quad i = 1..k \text{ e } l = 1..p \quad (3.26)$$

em que  $c_{il}$  e  $\sigma_{il}$  são os centros e as variâncias das várias funções de pertença.

As funções de pertença dos consequentes são funções constantes  $f_i = v_i$ , também conhecidas por *singletons* [Wan97].

O grau de pertença global de cada regra  $\beta_i$  é dado pelo produto algébrico (operador  $\prod$ ) dos graus de pertença dos conjuntos difusos correspondentes às várias entradas da referida regra, como definido na equação 3.27:

$$\beta_i = \prod_{l=1}^p \mu_{A_{il}}(x_l; c_{il}; \sigma_{il}), \quad i = 1..k \quad (3.27)$$

A operação normalização dos pesos (operador  $\mathbf{N}$ ) é realizada conforme definido na equação 3.28:

$$\gamma_i = \frac{\beta_i}{\sum_{i=1}^k \beta_i} = \frac{\prod_{l=1}^p e^{-\frac{(x_l - c_{il})^2}{2\sigma_{il}^2}}}{\sum_{i=1}^k \left( \prod_{l=1}^p e^{-\frac{(x_l - c_{il})^2}{2\sigma_{il}^2}} \right)}, \quad i = 1..k \quad (3.28)$$

Desta forma, aplicando o operador produto algébrico aos consequentes, a saída de cada regra  $y_i$  é o produto algébrico entre o grau de pertença normalizado  $\gamma_i$  e a constante  $v_i$  da respectiva regra, definido na equação 3.29:

$$\hat{y}_i = \gamma_i f_i = \gamma_i v_i, \quad i = 1..k \quad (3.29)$$

Finalmente, a saída global do sistema neuro-difuso  $\hat{y}$  é dada pela soma linear das saídas de cada regra  $\hat{y}_i$ , como apresentado na equação 3.30:

$$\hat{y} = \sum_{i=1}^k \hat{y}_i = \sum_{i=1}^k \gamma_i v_i = \sum_{i=1}^k \left( \frac{\prod_{l=1}^p e^{-\frac{(x_l - c_{il})^2}{2\sigma_{il}^2}}}{\sum_{i=1}^k \left( \prod_{l=1}^p e^{-\frac{(x_l - c_{il})^2}{2\sigma_{il}^2}} \right)} v_i \right), \quad i = 1..k \quad (3.30)$$

Com base num conjunto de dados de entrada e saída de dimensão  $N$ , o método do gradiente efectua a propagação do erro das saídas para as entradas de forma a minimizar uma dada função de custo  $J$ . Tipicamente, a função de custo a minimizar é dada por uma expressão semelhante à apresentada na equação 3.31:

$$J = \frac{1}{2} \sum_{n=1}^N e_n^2 = \frac{1}{2} \sum_{n=1}^N (y_n - \hat{y}_n)^2 \quad (3.31)$$

sendo  $y$  a saída do sistema e  $\hat{y}$  a saída estimada.

Com o objectivo de minimizar  $J$ , a regra do algoritmo de aprendizagem dos parâmetros  $v_i$ ,  $c_{il}$  e  $\sigma_{il}$  basear-se-á na regra da cadeia, utilizando a diferenciação parcial camada a camada até atingir os respectivos parâmetros.

Os parâmetros das funções de pertença dos consequentes  $v_i$  são identificados recorrendo aos respectivos Jacobianos definido na equação 3.32:

$$\frac{\partial J}{\partial v_i} = \frac{\partial J}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial \hat{y}_i} \frac{\partial \hat{y}_i}{\partial v_i} = -(y_n - \hat{y}_n) \frac{\beta_i}{\sum_{i=1}^k \beta_i} = -e\gamma_i, \quad i = 1..k \quad (3.32)$$

A regra iterativa de ajuste dos respectivos parâmetros é dada pela expressão definida na equação 3.33:

$$v_i(n+1) = v_i(n) + \lambda(n)e\gamma_i, \quad i = 1..k \quad (3.33)$$

em que  $\lambda(n)$  é a constante de aprendizagem (factor numérico que permite regular a velocidade de aprendizagem).

Os parâmetros dos antecedentes  $c_{il}$  e  $\sigma_{il}$ , centros e variâncias das respectivas funções de pertença, são identificados com base na regra da cadeia de diferenciação.

Para os parâmetros  $c_{il}$  os respectivos Jacobianos estão definidos nas equações 3.34 e 3.35:

$$\frac{\partial J}{\partial c_{il}} = \frac{\partial J}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial \hat{y}_i} \frac{\partial \hat{y}_i}{\partial \gamma_i} \frac{\partial \gamma_i}{\partial \beta_i} \frac{\partial \beta_i}{\partial \mu_{A_{il}}} \frac{\partial \mu_{A_{il}}}{\partial c_{il}}, \quad i = 1..k \text{ e } l = 1..p \quad (3.34)$$

$$\frac{\partial J}{\partial c_{il}} = -(y_n - \hat{y}_n) v_i \frac{\sum_{i=1}^k \beta_i - \beta_i}{\left( \sum_{i=1}^k \beta_i \right)^2} \frac{\beta_i}{\mu_{A_{il}}(x_l)} \left( \frac{1}{2} \frac{x_l - c_{il}}{\sigma_{il}^2} \mu_{A_{il}}(x_l) \right) \quad (3.35)$$

Depois de alguma manipulação algébrica da equação 3.35, os Jacobianos finais estão definidos na equação 3.36:

$$\frac{\partial J}{\partial c_{il}} = -\frac{1}{2} e (v_i - y) \gamma_i \frac{x_l - c_{il}}{\sigma_{il}^2} \quad (3.36)$$

A regra iterativa de ajuste dos parâmetros  $c_{il}$  é dada pela equação 3.37:

$$c_{il}(n+1) = c_{il}(n) + \lambda(n) \frac{1}{2} e \frac{x_l - c_{il}}{\sigma_{il}^2} (v_i - y), \quad i = 1..k \text{ e } l = 1..p \quad (3.37)$$

Para os parâmetros  $\sigma_{il}$  os respectivos Jacobianos estão definido na equação 3.38:

$$\frac{\partial J}{\partial \sigma_{il}} = \frac{\partial J}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial \hat{y}_i} \frac{\partial \hat{y}_i}{\partial \gamma_i} \frac{\partial \gamma_i}{\partial \beta_i} \frac{\partial \beta_i}{\partial \mu_{A_{il}}} \frac{\partial \mu_{A_{il}}}{\partial \sigma_{il}}, \quad i = 1..k \text{ e } l = 1..p \quad (3.38)$$

Deste modo, os Jacobianos finais são definidos pela equação 3.39:

$$\frac{\partial J}{\partial \sigma_{il}} = -e(v_i - y) \gamma_i \frac{(x_l - c_{il})^2}{\sigma_{il}^3} \quad (3.39)$$

Finalmente, a regra iterativa de ajuste dos parâmetros  $\sigma_{il}$  é dada pela equação 3.40:

$$\sigma_{il}(n+1) = \sigma_{il}(n) + \lambda(n)e(v_i - y) \gamma_i \frac{(x_l - c_{il})^2}{\sigma_{il}^3}, \quad i = 1..k \text{ e } l = 1..p \quad (3.40)$$

A execução iterativa destas regras de aprendizagem, com a correcta definição da constante de aprendizagem  $\lambda(n)$ , permite a identificação dos parâmetros internos do sistema neuro-difuso do tipo Takagi-Sugeno de ordem zero.

Este algoritmo iterativo deve ser executado enquanto o erro de treino não atingir um determinado valor predefinido, ou enquanto o erro de teste não atingir o seu valor mínimo. Se o erro de teste depois de atingir o mínimo começar a subir, o sistema neuro-difuso está a ajustar-se demasiado aos dados de treino perdendo capacidade de generalização (esta situação é vulgarmente designada por *overfitting* ou *overtraining*).

Uma das características deste algoritmo treino é poder “ficar preso” num mínimo local da função de custo, não atingindo o mínimo global. Existem técnicas que permitem aos algoritmos de aprendizagem sair de um mínimo local como, por exemplo, a técnica da variação da constante de aprendizagem  $\lambda(n)$ .

A aplicação do algoritmo de retropropagação do erro a outros tipos de sistemas neuro-difusos (com outras funções de pertença, outros operadores difusos ou outras funções nos consequentes) processa-se de forma semelhante à apresentada, com a excepção dos respectivos Jacobianos, que terão de ser redefinidos relativamente às novas funções de pertença, novos operadores difusos ou novas funções nos consequentes. Por exemplo, se as funções dos consequente fossem agora funções lineares relativamente às entradas em vez dos *singletons* utilizados, uma boa parte das derivadas parciais teriam de ser recalculadas contudo o procedimento seria semelhante ao apresentado.

### 3.5.2 Aprendizagem Híbrida

A aprendizagem híbrida de sistemas neuro-difusos consiste na utilização do método da retropropagação do erro, para o ajuste dos parâmetros não lineares das funções de pertença dos antecedentes, e na utilização do método de mínimos quadrados para determinar os parâmetros lineares dos consequentes.

Este método de aprendizagem realiza duas tarefas em cada iteração [Mat01]. Na primeira, os dados de entrada são propagados para a frente até à camada linear, sendo os parâmetros dos consequentes estimados pelo método de mínimos quadrados, enquanto os parâmetros dos antecedentes estão fixos; na segunda, o erro é retropropagado permitindo o ajuste dos parâmetros dos antecedentes, enquanto que os parâmetros dos consequentes das regras estão fixos. Este procedimento é repetido a cada iteração.

A tabela 3.2 sumaria os passos ou tarefas executadas a cada iteração.

Para demonstrar a aplicação deste algoritmo utiliza-se o exemplo ilustrativo apresentado na secção anterior de um sistema MISO neuro-difuso do tipo Takagi-Sugeno

	Propagação para a frente	Propagação para trás
Parâmetros Antecedentes	fixos	retropropagação do erro gradiente
Parâmetros Consequentes	estimador de mínimos quadrados	fixos

Tabela 3.2: Método híbrido de aprendizagem.

de ordem zero. Neste caso, a identificação dos parâmetros internos é realizada através da aplicação do algoritmo híbrido atrás descrito.

O sistema neuro-difuso em questão tem parâmetros lineares e parâmetros não lineares que serão identificados recorrendo à combinação dos dois algoritmos de aprendizagem, a retropropagação do erro e o estimador de mínimos quadrados.

A primeira parte do algoritmo híbrido consiste na identificação dos parâmetros lineares dos consequentes das regras utilizando o algoritmo de mínimos quadrados.

Uma vez que o vector das saídas do sistema neuro-difuso descrito na secção anterior é dado pela equação 3.41, a aplicação do algoritmo de mínimos quadrados torna-se muito simples:

$$\hat{y}_n = \gamma_{n1}v_1 + \gamma_{n2}v_2 + \dots + \gamma_{nk}v_k, \quad n = 1..N \quad (3.41)$$

ou seja

$$\hat{\mathbf{y}} = \mathbf{G}\mathbf{v} \quad (3.42)$$

Com o objectivo de calcular o vector dos parâmetros lineares dos consequentes  $\mathbf{v}$  (equação 3.43), é necessário propagar as entradas do sistema até à camada 4 da figura 3.10 e recolher os  $N$  vectores de dimensão  $k$  com os pesos normalizados para cada regra relativos às entradas (equação 3.44):

$$\mathbf{v} = [v_1 \ v_2 \ \dots \ v_k] \text{ vector coluna } k \times 1 \quad (3.43)$$

$$\mathbf{G} = [\gamma_{n1} \ \gamma_{n2} \ \dots \ \gamma_{nk}] \text{ matriz } N \times k \quad (3.44)$$

$$\mathbf{y} = [y_1 \ y_2 \ \dots \ y_N] \text{ vector coluna } N \times 1 \quad (3.45)$$

Finalmente, com a matriz  $\mathbf{G}$  e com o vector das saídas reais  $\mathbf{y}$  de dimensão  $N$  (equação 3.45), obtém-se o vector dos parâmetros dos consequentes  $\mathbf{v}$  de dimensão  $k$  definido na equação 3.46:

$$\mathbf{v} = [\mathbf{G}^T \mathbf{G}]^{-1} \mathbf{G}^T \mathbf{y} \text{ vector coluna } k \times 1 \quad (3.46)$$

A segunda parte do método híbrido consiste na identificação dos parâmetros não lineares dos antecedentes das regras  $c_{il}$  e  $\sigma_{il}$ . Estes parâmetros são identificados utilizando o algoritmo da retropropagação do erro apresentado na secção anterior.

Uma vez que os parâmetros não lineares são ajustados pelo método do gradiente de forma iterativa (lentamente), a determinação dos parâmetros lineares pelo método de mínimos quadrados não necessita de ser efectuada em todas as iterações. É típica a execução do algoritmo da retropropagação do erro em todas as iterações e o algoritmo de

mínimos quadrados só nas iterações múltiplas de dez (por exemplo). Este procedimento permite, mantendo a rapidez de convergência na identificação dos parâmetros, reduzir a carga computacional.

Para aumentar a velocidade de convergência deste algoritmo pode utilizar-se uma técnica que se baseia na variação da constante de aprendizagem  $\lambda(n)$ . Esta técnica consiste em incrementar a constante de aprendizagem  $\lambda(n)$  se, por exemplo, em cinco iterações seguidas o erro de modelização diminuir consecutivamente; por outro lado, esta técnica decrementa a constante de aprendizagem  $\lambda(n)$  se, por exemplo, em cinco iterações seguidas o erro de modelização aumentar consecutivamente.

O método híbrido é tipicamente utilizado em sistemas neuro-difusos do tipo Takagi-Sugeno, uma vez que estes possuem parâmetros lineares nos consequentes. A vantagem do método de identificação híbrido relativamente ao método totalmente baseado no gradiente está na velocidade de convergência. Pode-se salientar, também, que a utilização do método de mínimos quadrados garante a convergência dos respectivos parâmetros para um mínimo global.

Babuska propõe um método alternativo de identificação de sistemas neuro-difusos do tipo Takagi-Sugeno baseado no método de mínimos quadrados locais. Cada função dos consequentes de cada regra é identificada usando dados de entrada e saída próximos da zona de abrangência (locais) da referida regra. Esta condição permite a criação de modelos locais interpretáveis; no entanto, a utilização deste método provoca uma diminuição de precisão global do sistema [AB00].

### 3.6 Sistemas Neuro-Difusos Mais Comuns

Nas últimas décadas, a investigação na área da construção de aproximadores universais à custa de sistemas neuro-difusos registou um elevadíssimo número de contribuições, o que provocou o aparecimento de diversas variantes de sistemas neuro-difusos. Estas variantes são semelhantes na sua essência, contudo apresentam algumas diferenças, nomeadamente, ao nível da estrutura interna e dos algoritmos de identificação.

No final dos anos 80, início dos anos 90, surgiram os primeiros estudos na área das redes neuronais que aplicavam conceitos e ou operadores difusos, assim como na área dos sistemas difusos que aplicavam algoritmos de aprendizagem típicos da identificação das redes neuronais.

A introdução de conceitos e operadores difusos nas redes neuronais convencionais fez surgir as redes neuronais difusas [HBC92]. Um exemplo destas redes são as funções de base radial (*Radial Based Functions - RBF*) [JS93], que utilizam funções de pertença difusas a substituir as funções de activação. O objectivo desta junção foi o de obter redes neuronais adaptáveis com alguma interpretabilidade. Por outro lado, os sistemas interpretáveis difusos requeriam métodos de identificação automáticos, surgindo a ideia de utilizar os algoritmos típicos da identificação das redes neuronais nos sistemas difusos. Desde então, passou a ser possível ajustar os sistemas difusos de forma automática utilizando métodos de aprendizagem baseados no gradiente.

Os primeiros estudos dos sistemas neuro-difusos datam do início da década de 90, com Jang, Takagi, Hayashi, Lin e Lee em 1991; Berenji em 1992 e Nauck a partir



de 1993. A maioria das aplicações desta época eram voltadas para o controlo. Gradualmente foram aparecendo aplicações voltadas para outras áreas, como a análise e classificação de dados, detecção de falhas e sistemas de apoio à decisão.

Dos sistemas neuro-difusos mais conhecidos, destacam-se os seguintes:

- *Fuzzy Adaptive Learning Control Network* (FALCON) 1991, C. T. Lin e C. S. Lee [LL91];
- *Adaptive Network based Fuzzy Inference System* (ANFIS) 1992, R. R. Jang [Jan92] e [Jan93];
- *Generalized Approximate Reasoning based Intelligence Control* (GARIC) 1992, H. Berenji [BK92];
- *Neuronal Fuzzy Controller* (NEFCON) 1993, D. Nauck [Nau94];
- *Fuzzy Inference and Neural Network in Fuzzy Inference Software* (FINEST) 1996, Tano, O. e Arnould [TOA96];
- *Fuzzy Net* (FUN) 1993, S. Sulzberger, N. Tschichold e S. Vestli [STV93];
- *Artificial Neural Network Based Fuzzy Inference System* (ANNBFIS) 1997, J. Leski, E. Czogala [LC97] e [LC99];
- *Fuzzy Neural Network* (FNN) 1999, Figueiredo e Gomide [FG99];
- *Dynamic/Evolving Fuzzy Neural Network* (EFuNN and dmEFuNN) 1999, Kasabov e Song [KS99];
- *Self Constructing Neural Fuzzy Inference Network* (SONFIN) 1998, Juang e Lin [JL98].

Em [VDM04b] e [VM04d] são apresentados sumariamente os vários tipos de estruturas utilizadas nos sistemas neuro-difusos e os respectivos métodos de identificação.

É possível encontrar ainda em [Abr01] e [AN00] um estudo sobre o estado da arte deste tipo de sistemas e suas perspectivas de evolução. Nestes trabalhos são descritas várias estruturas neuro-difusas, métodos para a sua simplificação, algoritmos de identificação dos respectivos parâmetros e algumas aplicações. São ainda apresentadas as novas perspectivas de evolução e de simplificação destas estruturas baseadas em algoritmos genéticos.

## 3.7 Conclusão

Este capítulo apresentou vários métodos de construir e identificar modelos não lineares com base nos sistemas neuro-difusos. Foram apresentados vários tipos de modelos difusos construídos com base em conhecimentos adquiridos *a priori* e ajustados através

de métodos de identificação típicos das redes neuronais. Estes métodos de identificação permitem o ajuste automático destes modelos; no entanto, podem reduzir a sua interpretabilidade.

Na secção 3.2 apresentou-se uma descrição genérica dos sistemas difusos e dos seus blocos constituintes: fusificação, base de conhecimentos, inferência difusa e defusificação.

A secção 3.3 introduz os fundamentos e as estruturas das redes neuronais artificiais multi-camada, assim como os respectivos métodos de ajuste mais comuns.

Os sistemas neuro-difusos foram descritos na secção 3.4, sendo explorados dois sistemas em particular: sistemas neuro-difusos do tipo Mandani e do tipo Takagi-Sugeno. Esta secção apresenta ainda métodos para a escolha dos regressores mais significativos, para a definição do número/forma das funções de pertença e métodos para a optimização do número de regras difusas. Na secção 3.5 apresentaram-se e descreveram-se os dois métodos mais comuns de identificação destes modelos: o método baseado no gradiente e o método híbrido. Finalmente, na secção 3.6, apresentou-se um levantamento bibliográfico dos sistemas neuro-difusos mais comuns propostos na última década.

## Capítulo 4

# Modelização e Identificação Orientada por Blocos

### 4.1 Introdução

Os sistemas orientados por blocos, apresentados no capítulo 2, são classificados como métodos de modelização cinzenta [PP00]. Estes sistemas são constituídos pela combinação de duas funções ou blocos: um bloco não linear estático e um bloco linear dinâmico. A definição dos dois blocos e a forma de os combinar implica a existência e a utilização de conhecimentos adquiridos *a priori* construindo assim estruturas com um razoável nível de interpretabilidade.

Este capítulo descreve em detalhe as estruturas e os métodos de identificação de dois modelos orientados por blocos: Hammerstein e Wiener. As duas formas de combinar, em série, o bloco linear dinâmico com um bloco não linear estático dá origem aos dois tipos de modelos orientados por blocos já referidos. A origem da não linearidade estática é muitas vezes conhecida, podendo ter origem na não linearidade dos sensores ou dos actuadores ou ainda por uma relação física não linear existente entre as variáveis de entrada e saída do sistema como, por exemplo, a relação quadrática entre a corrente imposta numa resistência de aquecimento e a respectiva potência dissipada. Quando a origem da não linearidade estática é desconhecida esta é tipicamente modelizada por funções do tipo caixa cinzenta ou negra.

Na secção 4.2 são apresentadas e caracterizadas as estruturas dos modelos Hammerstein e Wiener. A secção 4.3 apresenta os diferentes métodos de identificação destes dois tipos de modelos incidindo, particularmente, nos vários métodos de identificação da função não linear estática. O tipo de dados disponíveis e o tipo de modelo orientado por blocos implicam a utilização de diferentes métodos de identificação tais como: a identificação em dois passos e a identificação iterativa ou num único passo [ABB<sup>+</sup>00].

### 4.2 Modelização Orientada por Blocos

Os efeitos não lineares encontrados em alguns sistemas industriais, como as colunas de destilação, sistemas de neutralização de pH, sistemas de transferência de calor ou

sistemas electromecânicos, podem ser modelizados eficientemente pela combinação de uma função estática não linear e uma função dinâmica linear [EJL91a] e [PP98].

Dependendo da forma de interligação dos blocos estático e dinâmico, definem-se vários métodos de modelização orientados por blocos como, por exemplo, [Abo03] e [PP00]:

- Modelização Hammerstein;
- Modelização Wiener.

A modelização da função linear dinâmica é tipicamente efectuada através de métodos lineares clássicos já bem conhecidos. A modelização da função não linear estática pode ser efectuada com base em conhecimentos *a priori* sobre o sistema uma vez que muitas vezes esta função pode ser observada e/ou é conhecida fenomenologicamente.

O comportamento não linear destes modelos pode ser removido do problema de identificação através da utilização de regressores não lineares, permitindo a utilização de algoritmos de aprendizagem lineares simples. Quando representados desta forma, os modelos Hammerstein e Wiener são casos particulares dos modelos NARX polinomiais [LB85], como será mostrado nas secções seguintes.

Na ausência de conhecimentos *a priori* sobre a função não linear estática, a sua modelização pode ser efectuada utilizando estruturas não lineares do tipo caixa cinzenta ou negra identificadas com base num conjunto de dados de entrada e saída. Neste caso, os métodos de identificação exigem uma carga computacional maior sendo um pouco mais morosos.

Além das formas apresentadas de interligar os dois blocos nos modelos Hammerstein e Wiener, apresentadas no capítulo 2, existem outras formas de o fazer [Coe02] como a apresentada no método de modelização orientado por blocos com realimentação [PP98], em que o bloco linear dinâmico é realimentado com o bloco não linear estático. De salientar o estudo comparativo efectuado por Pearson e Pottman [PP00] sobre a aplicação destes métodos de modelização a uma coluna de destilação.

O método de modelização orientado por blocos com realimentação é menos conhecido possuindo métodos de identificação mais complexos do que os utilizados nos outros dois métodos de modelização aqui referidos. Este método de modelização por blocos não será objecto de estudo no âmbito deste trabalho.

### 4.2.1 Modelização Linear

Os modelos Hammerstein e Wiener possuem um bloco linear dinâmico na sua estrutura interna. Um modelo linear dinâmico pode ser representado através da equação genérica 4.1:

$$A(q^{-1})y(k) = \frac{B(q^{-1})}{F(q^{-1})}u(k) + \frac{C(q^{-1})}{D(q^{-1})}\xi(k) \quad (4.1)$$

em que  $q^{-1}$  é o operador atraso unitário e  $A(q^{-1})$ ,  $B(q^{-1})$ ,  $C(q^{-1})$ ,  $D(q^{-1})$  e  $F(q^{-1})$  são funções polinomiais em  $q^{-1}$ ,  $y(k)$  é o vector das entradas,  $u(k)$  o é vector das saídas e

$\xi(k)$  é o vector que representa as perturbações aditivas do sistema, associado a sinais de natureza aleatória com média nula e uma dada variância (ruído branco).

Dependendo da definição de  $A(q^{-1})$ ,  $B(q^{-1})$ ,  $C(q^{-1})$ ,  $D(q^{-1})$  e  $F(q^{-1})$  as representações do respectivo modelo linear podem possuir várias designações como, por exemplo, ARMAX (*Auto Regressive Moving Average with eXogenous inputs*), ARX (*Auto Regressive with eXogenous inputs*), OE (*Output Error*), BJ (*Box Jenkins*).

Considerando um sistema linear SISO com uma entrada  $u(k)$ , uma saída  $y(k)$ , na presença de uma perturbação aditiva  $\xi(k)$ , a representação do sistema linear ARX obtém-se definindo  $F = C = D = 1$  na equação 4.1, o que resulta na equação 4.2:

$$A(q^{-1})y(k) = B(q^{-1})u(k) + \xi(k) \quad (4.2)$$

Baseado na equação 4.2 a estimação discreta da saída  $\hat{y}(k)$  é dada pela equação 4.3:

$$\hat{y}(k) = a_1 y(k-1) + \dots + a_{n_a} y(k-n_a) + b_1 u(k-d-1) + \dots + b_{n_b} u(k-d-n_b) \quad (4.3)$$

em que  $d$  é o tempo morto do sistema (múltiplo inteiro do período de amostragem  $h$ ),  $a_1, \dots, a_{n_a}$  e  $b_1, \dots, b_{n_b}$  são os parâmetros a identificar relativos às saídas e entradas passadas da função transferência definida na equação 4.2,  $n_a$  e  $n_b$  são, respectivamente, os números de regressores da saída e da entrada.

A estimação da saída  $\hat{y}(k)$ , expressa através da forma paramétrica da representação ARX, é dada pela equação 4.4:

$$\hat{y}(k) = \sum_{i=1}^{n_a} a_i y(k-i) + \sum_{j=1}^{n_b} b_j u(k-j-d) \quad (4.4)$$

Sabendo o vector dos parâmetros  $\theta$  (equação 4.5) e tendo o vector com os regressores  $\varphi$  (equação 4.6), é possível estimar a próxima saída do sistema sob a forma matricial, como definido na equação 4.7:

$$\theta = [a_1, a_2, \dots, a_{n_a}, b_1, \dots, b_{n_b}]^T \quad (4.5)$$

$$\varphi(k) = [y(k-1), \dots, y(k-n_a), u(k-d-1), \dots, u(k-d-n_b)]^T \quad (4.6)$$

$$\hat{y}(k) = \theta^T \varphi(k) \quad (4.7)$$

### 4.2.2 Modelização Hammerstein

A ligação em série ilustrada na figura 4.1 de uma função não linear estática  $f(\cdot)$  e de uma função linear dinâmica  $G(\cdot)$  é designada por modelo Hammerstein [NG66]:

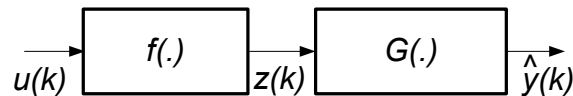


Figura 4.1: Modelo Hammerstein.

Pela figura 4.1, verifica-se que o sinal intermédio  $z(k)$  é obtido através da função estática  $f(\cdot)$ , como definido na equação 4.8:

$$z(k) = f(u(k)) \quad (4.8)$$

A estimação da saída  $\hat{y}(k)$  em função das amostras anteriores de  $z$  e da saída real  $y$  expressa através da representação de modelos ARX, está definida na equação 4.9:

$$\hat{y}(k) = \sum_{i=1}^{n_a} a_i \hat{y}(k-i) + \sum_{j=1}^{n_b} b_j z(k-d-j) \quad (4.9)$$

Substituindo a equação 4.8 em 4.9 obtém-se a estimação da saída através do modelo Hammerstein, onde a função estática não linear  $f(\cdot)$  é definida para todos os sinais de entrada atrasados, como definido na equação 4.10:

$$\hat{y}(k) = \sum_{i=1}^{n_a} a_i \hat{y}(k-i) + \sum_{j=1}^{n_b} b_j f(u(k-d-j)) \quad (4.10)$$

Pela equação 4.10 verifica-se que o modelo Hammerstein é um caso particular do modelo NARX polinomial [LB85], em que  $f(\cdot)$  é representada por um polinómio de grau  $l$ :

- a não linearidade está expressa só em função de  $u$ ;
- não existem termos do tipo  $u(k-d-i)u(k-d-j)$ , para  $i \neq j$ .

A saída do modelo em estado estacionário em resposta ao degrau  $\bar{y}$  pode ser expressa em função da entrada em estado estacionário  $\bar{u}$  como apresentada na equação 4.11:

$$\bar{y} = \frac{\left( \sum_{j=1}^{n_b} b_j \right) f(\bar{u})}{1 - \sum_{i=1}^{n_a} a_i} \quad (4.11)$$

em que  $\bar{u}$  e  $\bar{y}$  são as entradas e saídas em estado estacionário partindo do princípio que  $\bar{u} = u(k-d) = \dots = u(k-d-n_b)$  e que  $\bar{y} = y(k-1) = \dots = y(k-n_a)$ .

Pela equação 4.11 verifica-se que se  $f(\cdot)$  não for biunívoca os modelos Hammerstein podem apresentar vários valores de entrada  $\bar{u}$  que levam a uma mesma saída  $\bar{y}$ , ou seja, os modelos Hammerstein podem admitir multiplicidade dos valores de entrada [PP00].

A não linearidade estática representada nos modelos Hammerstein pode surgir das não linearidades introduzidas pelos sensores, actuadores ou das relações fenomenológicas verificadas entre as variáveis do sistema. A aplicação de dados de entrada e saída do sistema à equação 4.11 permite verificar a existência dessa função não linear estática, assim como a sua forma.

### 4.2.3 Modelização Wiener

A ligação em série ilustrada na figura 4.2 de uma função linear dinâmica  $G(\cdot)$  com uma função não linear estática  $f(\cdot)$  é designada por modelo Wiener [Wie58]:

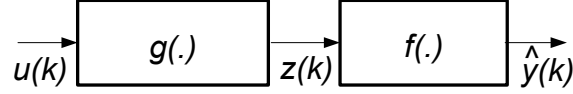


Figura 4.2: Modelo Wiener.

Pela figura 4.2, verifica-se que o sinal de saída  $\hat{y}(k)$  pode ser estimado pelo sinal intermédio  $z(k)$  através da função estática  $f(\cdot)$  como definido na equação 4.12:

$$\hat{y}(k) = f(z(k)) \quad (4.12)$$

Uma vez que, o sinal intermédio  $z(k)$  não pode ser medido, este pode ser calculado através da inversa da função  $f(\cdot)$ . Portanto, a função inversa de  $f(\cdot)$  tem de existir como definido na equação 4.13:

$$z(k) = f^{-1}(\hat{y}(k)) \quad (4.13)$$

O sinal intermédio  $z(k)$ , expresso através da representação de modelos ARX, está definido na equação 4.14:

$$z(k) = \sum_{i=1}^{n_a} a_i z(k-i) + \sum_{j=1}^{n_b} b_j u(k-d-j) \quad (4.14)$$

Substituindo a equação 4.14 na equação 4.12, obtém-se a estimação da saída do modelo Wiener, como definido na equação 4.15:

$$\hat{y}(k) = f \left( \sum_{i=1}^{n_a} a_i z(k-i) + \sum_{j=1}^{n_b} b_j u(k-d-j) \right) \quad (4.15)$$

Finalmente, substituindo a equação 4.13 na equação 4.15, a estimação da saída do modelo Wiener é obtida com base nas entradas e saídas anteriores do sistema, como definido na equação 4.16:

$$\hat{y}(k) = f \left( \sum_{i=1}^{n_a} a_i f^{-1}(\hat{y}(k-i)) + \sum_{j=1}^{n_b} b_j u(k-d-j) \right) \quad (4.16)$$

A saída em estado estacionário em resposta ao degrau  $\bar{y}$  pode ser expressa em função da entrada em estado estacionário  $\bar{u}$  como apresentado na equação 4.17:

$$\bar{y} = f \left( \frac{\left( \sum_{j=1}^{n_b} b_j \right) \bar{u}}{1 - \sum_{i=1}^{n_a} a_i} \right) \quad (4.17)$$

em que  $\bar{u}$  e  $\bar{y}$  são as entradas e saídas em estado estacionário partindo do pressuposto que  $\bar{u} = u(k-d) = \dots = u(k-d-n_b)$  e que  $\bar{y} = y(k-1) = \dots = y(k-n_a)$ .

Pela equação 4.17 verifica-se que se  $f(\cdot)$  não for biunívoca os modelos Wiener podem apresentar vários valores de entrada  $\bar{u}$  que levam a uma mesma saída  $\bar{y}$ , ou seja como os modelos Hammerstein, os modelos Wiener podem admitir multiplicidade dos valores de entrada [PP00].

A não linearidade estática representada nos modelos Wiener, como nos modelos Hammerstein, pode surgir das não linearidades introduzidas pelos sensores, actuadores ou das relações fenomenológicas existentes entre as entradas e saídas do sistema. A aplicação dos dados de entrada e saída do sistema à equação 4.17 permite verificar a existência dessa função não linear estática, assim como a sua forma.

### 4.3 Identificação de Modelos Orientados por Blocos

A maioria dos métodos de identificação de sistemas orientados por blocos são baseados nos modelos NARX polinomiais [LB85], [Cor01] e [Chi02]. Este modelos, além de exigirem um conjunto de dados de treino ricos em informação estática e dinâmica do sistema, podem tornar-se facilmente sobreparametrizados. Este problema pode originar a construção de modelos que podem exibir comportamentos locais incorrectos.

Os modelos orientados por blocos obedecem a um conjunto de restrições baseadas em conhecimentos adquiridos *a priori* que reduzem o problema apontado aos modelos NARX polinomiais. Essas restrições passam pela limitação do tipo de termos utilizados permitindo a construção de modelos não lineares mais simples. A utilização destes conhecimentos, particularmente relacionados com a função estática não linear, permitem definir e identificar modelos com um razoável grau de interpretabilidade (modelos cinzentos).

Dependendo do tipo de dados disponíveis e dos tipos de estruturas utilizadas para a construção dos dois tipos de modelos, os métodos de identificação dos modelos orientados por blocos podem dividir-se em duas classes:

- **Identificação em dois passos.** Esta classe é caracterizada por efectuar a identificação dos parâmetros da função linear dinâmica separadamente dos parâmetros da função não linear estática. Este procedimento requer a existência de dados de entrada e saída estáticos e dinâmicos.
- **Identificação iterativa ou num único passo.** Esta classe é caracterizada por ajustar todos os parâmetros da função linear dinâmica e da função não linear estática de forma iterativa [NG66] ou num único passo [KAWC95]. Este procedimento requer unicamente dados dinâmicos de entrada e saída.

As duas classes de identificação efectuem a identificação dos parâmetros fora de linha pressupondo que a dinâmica do sistema não se altera ao longo do tempo. Para sistemas cuja dinâmica se altera no tempo devem ser utilizadas técnicas de identificação linear em linha como, por exemplo, o método recursivo dos mínimos quadrados.

Como os modelos Hammerstein e Wiener são constituídos por dois blocos em série, o ganho do sistema pode ser modelizado por qualquer um destes blocos ou até pelos



dois. Isto pode levantar problemas de redundância na identificação dos dois blocos do modelo podendo até a identificação divergir [Sto81]. A imposição de algumas restrições na sua identificação pode evitar a referida redundância. No sentido de evitar o problema apontado é comum restringir o ganho da parte linear dinâmica a um ganho unitário [PP00], [ABB<sup>+</sup>00].

De referir que existem outros métodos de identificação que impõem outras restrições ao nível dos parâmetros do modelo linear dinâmico para evitar a redundância nos modelos criados, como é o caso, do método SVD (*Singular Value Decomposition*) [Bai98] e do método da pseudo inversa [BRD96]. Neste último método a restrição é imposta no parâmetro  $b_1$  do modelo linear dinâmico ( $b_1 = 1$ ), sendo a função não linear estática aproximada por uma função polinomial linear nos seus parâmetros. Deste modo é possível identificar os restantes parâmetros do modelo usando o método dos mínimos quadrados, eliminando a redundância nos modelos criados.

A classe de identificação em dois passos separados pode utilizar técnicas de identificação de parâmetros lineares ou não lineares. A identificação do bloco linear dinâmico é geralmente efectuada através da aplicação do método dos mínimos quadrados, aplicando ou não, algumas restrições de forma a não obter modelos redundantes. A identificação do bloco não linear estático depende do tipo de modelo utilizado, mas quando não linear utilizam-se normalmente os métodos baseados no gradiente ou em algoritmos genéticos.

A classe de identificação de um passo utiliza, tipicamente, o método dos mínimos quadrados. Neste caso, os modelos Hammerstein e Wiener incorporam a não linearidade estática ao nível dos seus regressores e não ao nível dos seus parâmetros, permitindo assim a sua identificação recorrendo a técnicas de regressão lineares.

### 4.3.1 Identificação do Bloco Linear Dinâmico Genérico

A identificação do bloco linear dinâmico  $G(\cdot)$  é realizada utilizando o método dos mínimos quadrados. Tendo disponíveis um conjunto de dados de entrada e saída  $\{u(k)\}$  e  $\{y(k)\}$ ,  $k = 1..N$ , e com o objectivo de minimizar a função de custo  $J$  definida na equação 4.18 (somatório dos quadrados da diferença entre a saída real  $y$  e a saída estimada  $\hat{y}$ ), o vector dos parâmetros lineares  $\boldsymbol{\theta} = [a_1, \dots, a_{n_a}, b_1, \dots, b_{n_b}]^T$  de dimensão  $(n_a + n_b)$  (equação 4.4) é estimado utilizando para isso a expressão genérica dos mínimos quadrados apresentada na equação 4.19:

$$J = \sum_{k=1}^N (y(k) - \hat{y}(k))^2 \quad (4.18)$$

$$\boldsymbol{\theta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (4.19)$$

em que  $\mathbf{y}$  é o vector com todas as saídas de dimensão  $N - k$  e  $\mathbf{X}$  é a matriz com todos os regressores de dimensão  $(N - k) \times (n_a + n_b)$ , como definido nas equações 4.20 e 4.21:

$$\mathbf{y} = [y(k+1), \dots, y(N)]^T \quad (4.20)$$

$$\mathbf{X} = \begin{bmatrix} y(k) & \dots & y(k - n_a) & u(k - d) & \dots & u(k - d - n_b) \\ y(k + 1) & & y(k + 1 - n_a) & u(k + 1 - d) & & u(k + 1 - d - n_b) \\ \dots & & \dots & \dots & & \dots \\ y(N - 1) & \dots & y(N - 1 - n_a) & u(N - 1 - d) & \dots & u(N - 1 - d - n_b) \end{bmatrix} \quad (4.21)$$

A identificação de modelos orientados por blocos equações 4.9 e 4.14 (bloco linear dinâmico) é sujeita à restrição de ganho unitário do bloco linear dinâmico. Assim, para a estimação do vector dos parâmetros  $\boldsymbol{\theta}$  impondo a restrição de ganho do bloco linear unitário, a condição definida na equação 4.22 tem de ser verificada:

$$\sum_{i=1}^{n_a} a_i + \sum_{j=1}^{n_b} b_j = 1 \text{ ou seja } \sum_{i=1}^{n_a+n_b} \tilde{\theta}_i = 1 \Leftrightarrow \mathbf{C}\tilde{\boldsymbol{\theta}} = s \quad (4.22)$$

em que  $\mathbf{C} = [1, 1, \dots, 1]$  de dimensão  $1 \times (n_a + n_b)$  e  $s$  é um escalar de valor unitário.

Draper e Smith [DS81] denominaram este método de minimização de  $J$  sujeito à restrição definida na equação 4.22 como, o método dos mínimos quadrados restrito (*constrained*). Designando  $\tilde{\boldsymbol{\theta}}$  o vector dos parâmetros calculados sujeitos à restrição imposta e  $\boldsymbol{\theta}$  o vector dos parâmetros sem restrição, então  $\tilde{\boldsymbol{\theta}}$  é dado pela equação 4.23:

$$\tilde{\boldsymbol{\theta}} = \boldsymbol{\theta} + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{C}^T [\mathbf{C}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{C}^T]^{-1} (\mathbf{C}\boldsymbol{\theta} - s) \quad (4.23)$$

Desta forma o modelo linear dinâmico identificado possui um ganho unitário evitando o problema da redundância.

### 4.3.2 Identificação dos Modelos Hammerstein

A identificação dos modelos Hammerstein é geralmente mais simples do que a identificação dos modelos Wiener, já que a função não linear só vai afectar os regressores do sinal de entrada. A forma como se define a função não linear  $f(\cdot)$  do modelo Hammerstein vai restringir o tipo de classe de identificação a utilizar.

A identificação do bloco linear dinâmico  $G(\cdot)$  é realizada utilizando o método dos mínimos quadrados que estima os parâmetros do modelo através de dados de entrada e saída dinâmicos do sistema, como apresentado na secção 4.3.1.

A função do bloco não linear estático  $f(\cdot)$  pode ser aproximada utilizando várias estruturas não lineares, tais como:

- funções matemáticas que descrevam o fenómeno;
- funções polinomiais;
- redes neuronais;
- funções de base radial (*Radial Based Functions - RBF*);
- séries de Volterra;
- sistemas difusos ou neuro-difusos;

Consoante a estrutura escolhida para o modelo Hammerstein (tipo de função que aproxima o bloco não linear) e o tipo de dados disponíveis do sistema, podem ser utilizadas duas classes de identificação.

### Classe de identificação em dois passos

A identificação dos modelos Hammerstein é efectuada, neste caso, em dois momentos separados. Se a função  $f(\cdot)$  for conhecida à partida através de conhecimentos adquiridos sobre o sistema ou conhecimentos fenomenológicos (primeiro passo), então o problema de identificação do modelo Hammerstein resume-se à identificação dos parâmetros do modelo linear dinâmico (segundo passo)  $\theta = [a_1, \dots, a_{n_a}, b_1, \dots, b_{n_b}]^T$  da equação 4.9 aplicando o método dos mínimos quadrados (equação 4.19) e utilizando os dados dinâmicos do sistema, em que  $\mathbf{y}$  é o vector com todas as saídas de dimensão  $N - k$  (equação 4.24) e  $\mathbf{X}$  é a matriz dos regressores (equação 4.25) em que  $z(k) = f(u(k))$ :

$$\mathbf{y} = [y(k+1), \dots, y(N)]^T \quad (4.24)$$

$$\mathbf{X} = \begin{bmatrix} y(k) & \dots & y(k-n_a) & z(k-d) & \dots & z(k-d-n_b) \\ y(k+1) & & y(k+1-n_a) & z(k+1-d) & & z(k+1-d-n_b) \\ \dots & & \dots & \dots & & \dots \\ y(N-1) & \dots & y(N-1-n_a) & z(N-1-d) & \dots & z(N-1-d-n_b) \end{bmatrix} \quad (4.25)$$

Quando a função  $f(\cdot)$  é desconhecida, existem como foi atrás referido, várias alternativas possíveis para a construção e identificação do seu modelo. Neste caso, são apresentadas duas abordagens para a modelização desta função que utilizam funções polinomiais e sistemas neuro-difusos.

Na primeira abordagem, a função  $f(\cdot)$  é modelizada por uma função polinomial de ordem  $l$  definida pela equação 4.26:

$$z(k) = f(u(k)) = o_1 u(k) + o_2 u(k)^2 + \dots + o_l u(k)^l \quad (4.26)$$

em que  $o_1, \dots, o_l$  são os parâmetros lineares do polinómio de grau  $l$  a identificar.

Na segunda abordagem, a função  $f(\cdot)$  é modelizada por um sistema difuso do tipo Takagi-Sugeno de ordem zero definido pela equação 4.27 (capítulo 3, secção 3.2):

$$z(k) = f(u(k)) = \frac{\sum_{i=1}^r \mu_{A_i}(u(k)) d_i}{\sum_{i=1}^r \mu_{A_i}(u(k))} \quad (4.27)$$

em que  $r$  é o número de regras do sistema difuso,  $\mu_{A_i}(u(k))$  é o grau de pertença relativo à entrada  $u(k)$  e ao conjunto difuso  $A_i$ , e em que  $d_i$  é a constante estática de saída da regra  $i$ . Para a definição dos conjuntos difusos utilizam-se funções de pertença do tipo Gaussiano (equação 4.28) cujo centro e a forma são definidos através dos parâmetros  $\sigma_i$  e  $c_i$  respectivamente. Neste caso, existem parâmetros lineares ( $d_i$ ) e não lineares ( $\sigma_i$  e  $c_i$ ). A função de pertença poderia ter outra forma: triangular, do tipo sino, ou outras,

variando número de parâmetros a identificar e a capacidade das respectivas estruturas na aproximação das não linearidades do sistema:

$$\mu_{A_i}(u(k)) = e^{-\frac{(u(k)-c_i)^2}{2\sigma_i^2}}, \quad i = 1..k \quad (4.28)$$

Se existirem dados de entrada e saída estáticos e dinâmicos disponíveis, então a função do bloco estático não linear é identificada (primeiro passo) seguida da identificação do bloco dinâmico linear (segundo passo). Se a função não linear  $f(\cdot)$  é modelizada por um polinómio de ordem  $l$  (equação 4.26) utilizando os dados estáticos de entrada e saída  $\{\bar{u}\}$  e  $\{\bar{z} = \bar{y}\}$ , os parâmetros  $o_1, o_2, \dots, o_l$  são estimados pelo método dos mínimos quadrados. Para evitar redundância nos modelos o ganho estático do sistema é todo incluído na função  $f(\cdot)$ , ficando a função do bloco linear dinâmico  $G(\cdot)$  com ganho unitário. A identificação dos parâmetros  $\theta$  do bloco linear é efectuada aplicando o método do mínimos quadrados (equação 4.19) definindo o vector  $\mathbf{y}$  com as saídas (equação 4.24) e a matriz  $\mathbf{X}$  dos regressores (equação 4.25), em que  $z(k) = f(u(k))$ .

Se a função  $f(\cdot)$  é modelizada por um sistema difuso (equação 4.27), então, utilizando conhecimentos experimentais e dados estáticos de entrada e saída ( $\{\bar{u}\}$  e  $\{\bar{z} = \bar{y}\}$ ), o conjunto de regras  $r$  é definido e os parâmetros  $\sigma_i$ ,  $c_i$  e  $d_i$  são estimados através dos métodos do gradiente ou híbrido, semelhantes aos apresentados no capítulo 3. Como já referido atrás, para evitar redundância nos modelos, o ganho estático do sistema é todo incluído na função  $f(\cdot)$ , ficando a função do bloco linear dinâmico  $G(\cdot)$  com ganho unitário. A identificação dos parâmetros  $\theta$  do bloco linear dinâmico é efectuada aplicando o método do mínimos quadrados (equação 4.19) definindo o vector  $\mathbf{y}$  com as saídas (equação 4.24) e a matriz  $\mathbf{X}$  dos regressores (equação 4.25) em que  $z(k) = f(u(k))$ .

Outra situação é quando só existem dados de entrada e saída dinâmicos disponíveis. Neste caso, os parâmetros do bloco dinâmico linear são identificados (primeiro passo), utilizando dados dinâmicos locais, através da expressão definida na equação 4.23 que impõe um ganho unitário, em que  $\mathbf{y}$  é o vector com todas as saídas de dimensão  $N - k$  e  $\mathbf{X}$  é a matriz com todos os regressores de dimensão  $(N - k) \times (n_a + n_b)$  definidos, respectivamente, nas equações 4.20 e 4.21.

De seguida são identificados os parâmetros da função estática não linear (segundo passo) utilizando dados dinâmicos globais do sistema. Utilizando as funções apresentadas para a modelização da função não linear estática (função polinomial e sistema neuro-difuso), a sua identificação é efectuada da forma já descrita, com a excepção de que dados utilizados são dados dinâmicos.

### Identificação iterativa ou num único passo

A identificação dos modelos Hammerstein é efectuada, neste caso, de uma forma iterativa ou num único passo. Desta classe de identificação destaca-se o método apresentado por Abonyi que identifica um modelo Hammerstein difuso utilizando programação quadrática (*quadratic programming*) [ABB<sup>+</sup>00] e o método apresentado por Narendra e Gallaman em que os dois blocos do modelo são identificados alternadamente de forma iterativa [NG66]. Este último método processa-se da seguinte forma: em cada iteração

numa primeira fase fixam-se os parâmetros  $a_i, b_j$  do bloco linear e identificam-se os parâmetros  $o_1, \dots, o_l$  da função não linear, através do método dos mínimos quadrados, numa segunda fase, com os parâmetros  $o_1, \dots, o_l$  identificados no passo anterior (primeira aproximação), identificam-se os parâmetros  $a_i, b_j$  do bloco linear (próxima aproximação), através do método dos mínimos quadrados. Iterativamente, este algoritmo irá convergir para a solução que minimiza a respectiva função de custo, no entanto, a respectiva solução pode apresentar redundância [Sto81]. Para evitar este problema utiliza-se a imposição de ganho unitário do bloco linear como definida nas equações 4.22 e 4.23.

Na classe de identificação Hammerstein, num único passo, destaca-se uma variante do método anterior com a imposição de restrições para evitar a redundância nos modelos em que todos os parâmetros são lineares e identificados de uma só vez pelo método dos mínimos quadrados.

### 4.3.3 Identificação dos Modelos Wiener

A identificação dos modelos Wiener é geralmente mais complexa do que a modelização Hammerstein já que a função não linear, equação 4.16, vai afectar os regressores do sinal de entrada e os regressores do sinal de saída.

A identificação do bloco linear dinâmico  $G(\cdot)$  é realizada utilizando o método dos mínimos quadrados. No entanto, a identificação do bloco linear dos modelos Wiener tem algumas particularidades que serão abordadas nesta secção.

O bloco não linear estático  $f(\cdot)$  pode ser modelizado utilizando várias estruturas não lineares, como as referidas para a modelização da função  $f(\cdot)$  nos modelos Hammerstein.

Consoante a estrutura escolhida para o modelo Wiener (tipo de função que aproxima o bloco não linear) e o tipo de dados disponíveis do sistema podem ser utilizadas duas classes de identificação.

#### Classe de identificação em dois passos

A identificação dos modelos Wiener é efectuada, neste caso, em dois momentos separados. Se a função  $f(\cdot)$  for conhecida à partida, então o problema de identificação do modelo Wiener resume-se a identificar os parâmetros lineares  $\theta = [a_1, \dots, a_{n_a}, b_1, \dots, b_{n_b}]^T$  através do método dos mínimos quadrados (equação 4.19) utilizando os dados dinâmicos do sistema (equação 4.14), em que  $\mathbf{X}$  é a matriz dos regressores definida na equação 4.29 com  $z(k) = f^{-1}(y(k))$ ,  $\mathbf{y}$  é o vector das saídas de dimensão  $N - k$  definido na equação 4.30 e  $f(\cdot)$  é uma função conhecida e que admite inversa  $f^{-1}(\cdot)$ :

$$\mathbf{X} = \begin{bmatrix} z(k) & \dots & z(k - n_a) & u(k - d) & \dots & u(k - d - n_b) \\ z(k + 1) & & z(k + 1 - n_a) & u(k + 1 - d) & & u(k + 1 - d - n_b) \\ \dots & & \dots & \dots & & \dots \\ z(N - 1) & \dots & z(N - 1 - n_a) & u(N - 1 - d) & \dots & u(N - 1 - d - n_b) \end{bmatrix} \quad (4.29)$$

$$\mathbf{y} = [z(k + 1), \dots, z(N)]^T = [f^{-1}(y(k + 1)), \dots, f^{-1}(y(N))]^T \quad (4.30)$$

Como a utilização do método dos mínimos quadrados penaliza o erro entre o valor de  $z(k)$  calculado e estimado, e não entre a saída  $y(k)$  real e estimada, por vezes o modelo criado aproxima correctamente  $z(k)$  mas também incorrectamente  $y(k)$  [PP00]. Para ultrapassar este problema podem ser utilizadas duas soluções. A primeira é evitar a transformação  $f(\cdot)$  e considerar o modelo Wiener como um problema de regressão de mínimos quadrados não linear; a segunda, utilizar a regressão de mínimos quadrados linear pesada, apresentada em [DS81], cuja função de custo é função do erro quadrático pesado. O peso introduzido na função de custo é função do quadrado da derivada de  $z(k)$ .

Quando a função  $f(\cdot)$  é desconhecida existem várias alternativas possíveis para a construção e identificação do seu modelo. Utilizando as mesmas abordagens referidas para a modelização da função não linear estática nos modelos Hammerstein esta função pode ser modelizada por funções polinomiais e sistemas neuro-difusos.

Se existem dados de entrada e saída estáticos e dinâmicos, a função do bloco estático não linear é identificada (primeiro passo) seguida da identificação do bloco dinâmico linear (segundo passo). Se a função  $f(\cdot)$  é uma função polinomial de ordem  $l$  (equação 4.26), utilizando os dados estáticos de entrada e saída ( $\{\bar{u} = \bar{z}\}$  e  $\{\bar{y}\}$ ), os parâmetros  $o_1, o_2, \dots, o_l$  são estimados pelo método dos mínimos quadrados. Para evitar redundância nos modelos o ganho estático do sistema é todo incluído na função  $f(\cdot)$ , ficando a função do bloco linear dinâmico  $G(\cdot)$  com ganho unitário. A identificação dos parâmetros  $\theta$  do bloco linear é efectuada aplicando o método do mínimos quadrados (equação 4.19) definindo o vector  $\mathbf{y}$  com as saídas (equação 4.30) e a matriz  $\mathbf{X}$  dos regressores (equação 4.29) em que  $z(k) = f^{-1}(y(k))$ , sendo  $f^{-1}(\cdot)$  é a função inversa de  $f(\cdot)$ .

Se a função  $f(\cdot)$  é aproximada por um sistema difuso (equação 4.27), eis que, utilizando conhecimentos experimentais e dados estáticos de entrada e saída ( $\{\bar{u}\}$  e  $\{\bar{z} = \bar{y}\}$ ) o conjunto de regras  $r$  é definido e os parâmetros  $\sigma_i$ ,  $c_i$  e  $d_i$  são estimados através dos métodos do gradiente ou híbrido, apresentados no capítulo 3. O ganho estático do sistema é todo incluído na função  $f(\cdot)$  evitando a redundância de modelos. A identificação dos parâmetros  $\theta$  do bloco linear é efectuada aplicando o método do mínimos quadrados (equação 4.19) definindo o vector  $\mathbf{y}$  com as saídas (equação 4.30) e a matriz  $\mathbf{X}$  dos regressores (equação 4.29) em que  $z(k) = f^{-1}(y(k))$ .

De notar que inverter, matematicamente, um sistema difuso não é tarefa fácil, sendo a maioria das vezes até impossível [Sou03]. A grande diferença relativamente aos modelos Hammerstein é que nos modelos Wiener é necessário conhecer as funções inversas  $f^{-1}(\cdot)$  das funções  $f(\cdot)$  definidas nas equações 4.26 e 4.27, para poder estimar o vector dos parâmetros  $\theta$  do bloco linear através do método dos mínimos quadrados. Se a função  $f(\cdot)$  admitir inversa  $f^{-1}(\cdot)$ , utiliza-se essa inversa; se a função inversa for difícil ou mesmo impossível de obter a partir da função  $f(\cdot)$ , pode recorrer-se à identificação de uma aproximação da função inversa  $f^{-1}(\cdot)$  designada por  $\hat{f}^{-1}(\cdot)$ , utilizando uma função polinomial de ordem  $l$  ou um sistema neuro-difuso (equações 4.31 e 4.32, respectivamente):

$$z(k) = \hat{f}^{-1}(y(k)) = s_1 y(k) + s_2 y(k)^2 + \dots + s_l y(k)^l \quad (4.31)$$

$$z(k) = \hat{f}^{-1}(y(k)) = \frac{\sum_{i=1}^r \mu_{A_i}(y(k))v_i}{\sum_{i=1}^r \mu_{A_i}(y(k))} \quad (4.32)$$

Caso só existam dados de entrada e saída dinâmicos disponíveis, o processo de identificação é semelhante ao descrito para o caso da identificação dos modelos Hammerstein em dois passos.

### Identificação iterativa ou num único passo

A identificação de todos os parâmetros dos modelos Wiener é efectuada de forma iterativa ou num único passo. Da classe de identificação de modelos Wiener iterativa destaca-se o método apresentado por Wigren [Wig93]. Este método é semelhante ao apresentado por Narendra e Gallaman em [NG66] para a identificação de modelos Hammerstein iterativo. Baseado no método anterior, Kalafatis [KAWC95] apresenta a identificação efectuada num único passo impondo a restrição  $b_1 = 1$  ao parâmetro do modelo linear dinâmico evitando a identificação de modelos redundantes. O autor calcula os parâmetros do bloco linear e os da função inversa  $f^{-1}(\cdot)$  simultaneamente num único passo. Só depois deste procedimento é que calcula a função  $f(\cdot)$ . Neste caso, a função inversa  $f^{-1}(\cdot)$  e  $f(\cdot)$  são ambas aproximadas por uma função polinomial de grau baixo.

Outras alternativas para a identificação de modelos Wiener passam pela utilização de programação evolutiva (*Evolutionary Programing - EP*) [HUK03] ou algoritmos da máxima semelhança (*Maximum Likelihood - ML*) [Lju00]. Estes métodos de identificação são computacionalmente mais exigentes do que os métodos atrás referidos.

## 4.4 Conclusão

Este capítulo apresentou dois métodos de modelização orientados por blocos designados por modelização Hammerstein e Wiener. Estes métodos permitem a modelização de um subgrupo de sistemas não lineares de uma forma relativamente simples. Basicamente estes métodos subdividem a função não linear global do sistema em dois blocos: um bloco linear dinâmico e outro não linear estático. A forma como se conjugam os blocos, a definição dos regressores e a aplicação de conhecimentos fenomenológicos ou experimentais na definição da função não linear estática, são formas de incorporar um determinado nível de interpretabilidade na construção destes modelos. O facto destes modelos possuírem alguma interpretabilidade permite a sua classificação em modelos do tipo caixa cinzenta.

Na secção 4.2 apresentou-se uma descrição dos modelos orientados por blocos mais conhecidos. Foram definidas e apresentadas as funções matemáticas dinâmicas e estáticas dos modelos Hammerstein e Wiener.

Foram ainda apresentadas, na secção 4.3, várias técnicas de identificação para os modelos Hammerstein e Wiener. Estas técnicas dividiram-se fundamentalmente em duas classes: identificação em dois passos e identificação iterativa ou num único passo.

A classe de identificação utilizada depende, principalmente, do tipo de dados disponíveis e do tipo de função escolhida para a modelização do bloco não linear estático.

Este tipo de modelos possibilita a incorporação de conhecimentos *a priori* sobre o sistema de uma forma simples, permitindo a modelização de um subgrupo de sistemas não lineares à custa da modelização linear ou à custa da conjugação desta com modelos não lineares relativamente simples.



## Capítulo 5

# Modelização e Identificação Híbrida

### 5.1 Introdução

A combinação de um modelo fenomenológico com um modelo desconhecido é conhecida por modelo híbrido, também designado *semi-mechanistic* [ARBS03]. Estes são construídos através da combinação em série, ou em paralelo, de um modelo fenomenológico aproximado (do tipo caixa branca), com um modelo desconhecido (do tipo caixa cinza escura ou negra).

A modelização híbrida referida no capítulo 2 foi apresentada como um método de modelização do tipo cinza claro. Esta classificação é resultado do tipo de conhecimentos utilizados na construção deste tipo de estruturas. Os modelos híbridos baseiam-se, essencialmente, nos princípios fenomenológicos do sistema permitindo a construção de estruturas com um elevado grau de interpretabilidade e uma elevada capacidade de generalização.

Fundamentalmente, o modelo híbrido baseia-se nos primeiros princípios, ou seja, nas equações de balanço do sistema. No entanto, para sistemas com alguma complexidade existem partes destas equações que são mal conhecidas ou até desconhecidas. Essas partes desconhecidas são modelizadas por um segundo bloco do tipo caixa cinza escura ou negra como, por exemplo, uma rede neuronal ou um sistema neuro-difuso.

Deste modo, esta metodologia de modelização maximiza a utilização de conhecimentos *a priori* sobre o sistema baseando-se, principalmente, em conhecimentos fenomenológicos complementados por conhecimentos experimentais e dados de entrada e saída.

A identificação deste tipo de modelos é efectuada em duas fases: definição do modelo fenomenológico e identificação do modelo desconhecido. A definição do modelo fenomenológico e dos seus parâmetros internos é, geralmente, feita à custa das equações diferenciais do sistema e de algumas medidas estáticas efectuadas no sistema para validar os respectivos parâmetros internos. O modelo do tipo caixa cinzenta ou negra que modeliza a parte desconhecida é identificado com base num conjunto de dados de entrada e saída do sistema e com base no respectivo modelo fenomenológico.

Este capítulo, na secção 5.2, apresenta a solução de modelos híbridos para a modelização de sistemas não lineares. É apresentado e definido um modelo fenomenológico genérico, a estrutura do modelo híbrido série e, de seguida, a estrutura do modelo

híbrido paralelo.

A identificação dos parâmetros de um modelo híbrido surge na secção 5.3. Neste ponto são apresentados e discutidos os métodos utilizados para o cálculo dos parâmetros do modelo fenomenológico e os métodos de identificação dos parâmetros do modelo desconhecido nos modelos híbridos série e paralelo.

## 5.2 Modelização Híbrida

O conhecimento fenomenológico dos sistemas é muito importante para a construção de modelos, principalmente para a construção de modelos de sistemas complexos não lineares. Quanto mais profundos forem os conhecimentos acerca dos princípios fenomenológicos, melhor será o modelo fenomenológico. No entanto, nem sempre são bem conhecidas todas as relações fenomenológicas como, por exemplo, a cinemática das reacções em processos biológicos [Oli04] ou a cinemática das reacções de aquecimento/arrefecimento em processos térmicos [XJ02]. Para solucionar este problema, surgem os sistemas híbridos que permitem a criação de modelos baseados na ligação em série, ou paralelo, de um modelo fenomenológico com um modelo desconhecido. O modelo desconhecido é, geralmente, um bloco não linear do tipo caixa negra ou cinzenta que modeliza as partes das relações fenomenológicas menos conhecidas ou até desconhecidas do sistema real. Assim, a combinação dos dois blocos permite a criação de modelos globais, não só com elevada interpretabilidade e capacidade de generalização, mas também com reduzida complexidade matemática.

Estes métodos de modelização são muito utilizados na criação de modelos de sistemas altamente não lineares como processos químicos [TK94a], [CAPL96], [PU92b], [ZSK96], processos biológicos [TR96], [SL97], [CBH<sup>+</sup>96] e processos térmicos [KFT03], [VM04f].

O modo de combinar o modelo desconhecido com o modelo fenomenológico pode ser efectuado, tipicamente, de duas formas: em série ou em paralelo, dando origem aos dois modelos com características diferentes [PU92a], [TKB92]:

- Modelo híbrido série;
- Modelo híbrido paralelo.

A estrutura dos modelos híbridos é fortemente dependente do sistema real em causa, uma vez que estes se baseiam nas equações de balanço macroscópicas dos sistemas em causa como, por exemplo, balanço de massa, de momento ou de energia. Contudo, é possível definir uma equação de balanço genérica adaptável a qualquer sistema. Com base nesta equação genérica e nos conhecimentos fenomenológicos existentes sobre o sistema em causa, a definição da equação de balanço específica é efectuada com relativa facilidade.

O modelo desconhecido a combinar com o modelo fenomenológico pode ser construído à custa de qualquer método de modelização não linear. No entanto, utilizam-se normalmente métodos de modelização do tipo caixa cinzenta escura ou negra como, as redes neuronais [TK94a], [Oli04] e os sistemas neuro-difusos [RKS<sup>+</sup>99], [KFT03].

### 5.2.1 Definição do Modelo Fenomenológico

O primeiro passo, e o mais importante, na definição e identificação de um modelo híbrido é a definição do modelo fenomenológico. Geralmente, os modelos fenomenológicos, modelos do tipo caixa branca (figura 5.1), são definidos a partir das equações macroscópicas do sistema tais como o balanço de momento, massa ou energia [Fra88], [Bej93], [ARBS03]:

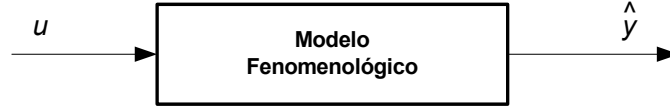


Figura 5.1: Modelo fenomenológico.

A equação 5.1 apresenta o balanço genérico baseado nos princípios da conservação de energia que levam à escrita das equações diferenciais genéricas, onde  $x_i$  é uma determinada grandeza, como o momento, a massa ou a energia de um dado sistema:

$$\begin{bmatrix} \text{variação} \\ \text{de } x_i \end{bmatrix} = \begin{bmatrix} \text{fluxo de} \\ \text{entrada de } x_i \end{bmatrix} - \begin{bmatrix} \text{fluxo de} \\ \text{saída de } x_i \end{bmatrix} + \begin{bmatrix} \text{quantidade} \\ \text{de } x_i \text{ gerado} \end{bmatrix} - \begin{bmatrix} \text{quantidade de} \\ x_i \text{ consumido} \end{bmatrix} \quad (5.1)$$

A equação 5.1 pode ser formulada utilizando a representação no espaço de estados, resultando na equação 5.2:

$$\begin{aligned} \dot{\mathbf{x}} &= f_{PP}(\mathbf{x}, \mathbf{u}) \\ \hat{\mathbf{y}} &= g(\mathbf{x}) \end{aligned} \quad (5.2)$$

onde  $\mathbf{x} = [x_1, x_2, \dots, x_m]^T$  representa o vector de estados de dimensão  $m$ ,  $\hat{\mathbf{y}} = [\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n]^T$  representa o vector da estimação das saídas de dimensão  $n$  e  $\mathbf{u} = [u_1, u_2, \dots, u_n]^T$  representa o vector das entradas do sistema de dimensão  $n$ . Em geral,  $f_{PP}(\cdot)$  e  $g(\cdot)$  são funções não lineares, em que  $f_{PP}(\cdot)$  representa a função dos primeiros princípios do sistema.

Os modelos do tipo caixa branca são validados por experiências que permitem a confirmação dos seus parâmetros internos. Na prática, a construção de modelos fenomenológicos pode-se tornar numa tarefa difícil, uma vez que, nem sempre as relações fenomenológicas subjacentes são completamente conhecidas, podendo mesmo existir relações no modelo completamente desconhecidas.

Geralmente, nem todos os termos da equação 5.1 são completamente conhecidos. Assim, quando não existe conhecimento de todos os detalhes dos primeiros princípios, partes do modelo global são modelizadas através de modelos ou funções do tipo caixa cinzenta escura ou negra  $f_{CCN}(\mathbf{x}, \mathbf{u}, \boldsymbol{\theta})$ .

O modelo híbrido pode ser formulado através da equação 5.3, baseada na equação de balanço genérica (equação 5.2), incluindo internamente a parte desconhecida modelizada por uma função do tipo caixa cinzenta escura ou negra:

$$\begin{aligned}\dot{\mathbf{x}} &= f_{PP}(\mathbf{x}, \mathbf{u}, f_{CCN}(\mathbf{x}, \mathbf{u}, \boldsymbol{\theta})) \\ \hat{\mathbf{y}} &= g(\mathbf{x})\end{aligned}\tag{5.3}$$

em que  $f_{PP}(\cdot)$ <sup>1</sup> representa a função dos primeiros princípios (equações de balanço),  $g(\cdot)$  é, genericamente, uma função não linear, e  $f_{CCN}(\cdot)$  representa a função desconhecida modelizada por um modelo do tipo caixa cinzenta ou negra parametrizada através do vector  $\boldsymbol{\theta}$ .

As taxas de reacção nos sistemas químicos ou biológicos (relação entre a quantidade de  $x_i$  gerado e consumido) e processos de transferência de calor (relação entre o calor fornecido e libertado) são fenómenos particularmente difíceis de definir através de equações físicas com parâmetros conhecidos. No entanto, os primeiros dois termos de transporte (relação entre fluxos) da equação 5.1 são fenómenos descritos facilmente e de uma forma exacta. Assim, na fase de modelização, é perceptível qual a parte do modelo fenomenológico mais fácil de obter e qual a parte mais trabalhosa.

Um caso particular do modelo definido na equação 5.3 é o modelo definido na equação 5.4, em que a variação da grandeza momento, massa ou energia em ordem ao tempo é igual à soma da função conhecida  $f_{PP}(\mathbf{x}, \mathbf{u})$  com a função  $f_{CCN}(\mathbf{x}, \mathbf{u}, \boldsymbol{\theta})$ :

$$\begin{aligned}\dot{\mathbf{x}} &= f_{PP}(\mathbf{x}, \mathbf{u}) + f_{CCN}(\mathbf{x}, \mathbf{u}, \boldsymbol{\theta}) \\ \hat{\mathbf{y}} &= g(\mathbf{x})\end{aligned}\tag{5.4}$$

onde  $f_{PP}(\cdot)$  representa a função dos primeiros princípios,  $g(\cdot)$  é uma função não linear e  $f_{CCN}(\mathbf{x}, \mathbf{u}, \boldsymbol{\theta})$  representa a função difícil, ou até impossível, de definir através das equações diferenciais como, por exemplo, o balanço entre a energia fornecida e libertada para o exterior num sistema térmico.

### 5.2.2 Modelo Híbrido Série

O modelo híbrido série consiste na combinação em série, tal como representado na figura 5.2, do modelo fenomenológico com o modelo do tipo caixa cinzenta ou negra [PU92a]. Na figura,  $\mathbf{u}$  representa a entrada,  $\hat{\mathbf{y}}$  representa a estimação da saída e  $\mathbf{z}$  representa a saída do modelo do tipo caixa cinzenta ou negra definido por  $f_{CCN}(\mathbf{x}, \mathbf{u}, \boldsymbol{\theta})$ :

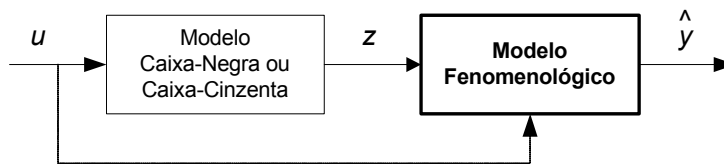


Figura 5.2: Modelo híbrido série.

<sup>1</sup>Esta função  $f_{PP}(\cdot)$  não é igual à função referida na equação 5.2, mas para não aumentar o número de símbolos será designada pelo mesmo nome.

A modelização híbrida série é utilizada quando existem conhecimentos sobre a existência da função mal conhecida, ou desconhecida, que, na impossibilidade de a definir através das equações diferenciais, tem de ser modelizada por outro método. Nesta situação, a parte desconhecida é modelizada utilizando uma função do modelo do tipo caixa cinzenta ou negra cujos respectivos regressores são definidos com base em conhecimentos fenomenológicos e/ou experimentais existentes sobre essa parte desconhecida. Portanto, a modelização híbrida série deve ser utilizada quando existem conhecimentos *a priori* sobre a parte desconhecida, ou mal conhecida, e que fenómeno é que essa parte representa no modelo fenomenológico global.

O primeiro passo para a definição do modelo híbrido série é a definição do modelo fenomenológico do sistema, como apresentado na secção 5.2.1. O segundo, é a definição do modelo do tipo caixa cinzenta ou negra que vai modelizar a parte mal conhecida ou até desconhecida do sistema.

O modelo do tipo caixa cinzenta ou negra pode ser construído utilizando qualquer tipo de modelização não linear. No entanto, como a função a modelizar é mal conhecida, ou desconhecida, é típica a utilização de redes neuronais ou de sistemas neuro-difusos, uma vez que basta um conjunto de dados de entrada e saída para a sua identificação.

As redes neuronais artificiais, perceptrões de multi-camada (MLP) ou funções de base radial (RBF) são, por dois motivos, as funções mais utilizadas para modelização da parte desconhecida do sistema. Primeiro, porque estas funções permitem a modelização de qualquer função linear ou não linear; segundo, porque a sua aplicação não requer praticamente nenhum conhecimento *a priori* sobre o sistema a modelizar.

Aplicações de redes neuronais na modelização da parte desconhecida do modelo híbrido série podem ser encontradas em [Oli04], [PU92a] e [NH04].

Os sistemas neuro-difusos passaram também a ser utilizados mais recentemente por dois motivos. Primeiro, porque permitem a aproximação de qualquer função linear ou não linear; segundo, porque com alguns conhecimentos *a priori* sobre a parte menos conhecida do sistema é possível criar um modelo da função desconhecida com algum nível de interpretabilidade. Esta abordagem permite a criação de modelos híbridos mais transparentes e interpretáveis.

Aplicações de sistemas neuro-difusos na modelização da parte desconhecida do modelo híbrido série podem ser encontradas em [RKS<sup>+</sup>99], [ARBS03], [VM04f] e [VM05].

### 5.2.3 Modelo Híbrido Paralelo

O modelo híbrido paralelo consiste na combinação em paralelo, tal como representado na figura 5.3, do modelo fenomenológico aproximado com o modelo do tipo caixa cinzenta ou negra [TKB92]. Na figura  $\hat{y}$  representa a saída do modelo fenomenológico aproximado,  $\hat{y}_d$  a saída do modelo desconhecido, e  $\hat{y}_g$  é a saída global.

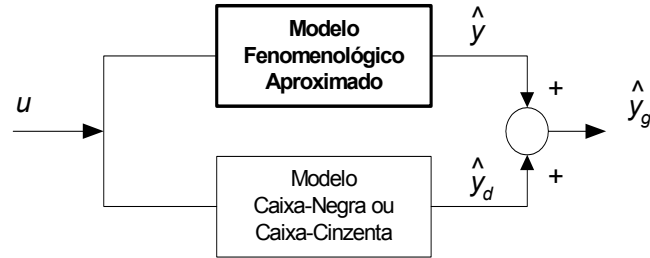


Figura 5.3: Modelo híbrido paralelo.

A saída  $\hat{y}_g$  é a soma das saídas de dois modelos: do modelo fenomenológico aproximado  $\hat{y}$  e do bloco do tipo caixa cinzenta ou negra  $\hat{y}_d$ .

A modelização híbrida paralela utiliza-se, normalmente, quando não existem nenhuns conhecimentos *a priori* sobre a origem fenomenológica do erro existente entre o modelo fenomenológico aproximado e o sistema real. Neste caso, não se trata de modelizar a parte mal conhecida do sistema em que se conhecem as variáveis que estão envolvidas e o fenómeno em causa (modelização híbrida série). Trata-se sim de modelizar a diferença entre o sistema real e o modelo fenomenológico aproximado cuja origem é completamente desconhecida.

Tal como na modelização híbrida série, o primeiro passo para a definição do modelo híbrido paralelo é a definição do modelo fenomenológico do sistema, como apresentado na secção 5.2.1. O segundo passo é a definição do modelo do tipo caixa cinzenta ou negra que vai modelizar a diferença entre a saída do modelo fenomenológico e do sistema real.

O modelo fenomenológico é definido através das equações de balanço do sistema (equação 5.2); no entanto, verifica-se que a saída deste não representa correctamente o sistema real. A origem desta diferença é completamente desconhecida. Assim, a estratégia para a sua modelização passa pela utilização de um modelo do tipo caixa cinzenta ou negra  $f_{CCN}(\mathbf{x}, \mathbf{u}, \boldsymbol{\theta})$  colocado em paralelo com o modelo fenomenológico aproximado.

O modelo do tipo caixa cinzenta ou negra  $f_{CCN}(\mathbf{x}, \mathbf{u}, \boldsymbol{\theta})$  a utilizar é construído à custa de redes neuronais ou de sistemas neuro-difusos. Estes métodos de modelização permitem a identificação dos respectivos modelos sem utilizar conhecimentos *a priori*, recorrendo simplesmente a um conjunto de dados de entrada e saída.

Assim, e pelos mesmos motivos enunciados na escolha do método de modelização do modelo desconhecido na modelização híbrida série, as redes neuronais artificiais são muito utilizadas, como efectuado em [XJ02], [NH04] e [TZM01]. Em [TZM01] são utilizadas três redes neuronais para a estimação de três variáveis do sistema, sendo as respectivas saídas somadas à saída do modelo fenomenológico aproximado.

Os sistemas neuro-difusos também podem ser utilizados para a modelização da parte desconhecida [KFT03]. Neste artigo, o autor utiliza um modelo fenomenológico aproximado colocando em paralelo com um sistema neuro-difuso, cuja soma das saídas dos respectivos modelos é o valor de saída do modelo híbrido paralelo.

## 5.3 Identificação de Modelos Híbridos

A identificação dos modelos híbridos consiste na determinação dos parâmetros do modelo fenomenológico e na identificação dos parâmetros do modelo do tipo caixa cinzenta ou negra. Deste modo, a identificação dos modelos híbridos subdivide-se em duas partes:

- **Determinação dos parâmetros do modelo fenomenológico.** A determinação dos parâmetros do modelo fenomenológico é realizada através da execução de alguns testes experimentais. Os conhecimentos fenomenológicos e os testes realizados no sistema permitem a obtenção dos parâmetros estáticos necessários para a definição do modelo fenomenológico.
- **Identificação dos parâmetros do modelo mal conhecido ou desconhecido.** A identificação do modelo desconhecido depende do tipo de modelo (cinzento ou negro) escolhido. Genericamente utilizam-se métodos de modelização do tipo caixa cinzenta ou negra cuja respectiva identificação necessita de um conjunto de dados de entrada e saída. A parte desconhecida é tipicamente modelizada por uma função não linear e identificada recorrendo a algoritmos de aprendizagem iterativos baseados na derivada do erro. Contudo, se a parte desconhecida for modelizada por funções lineares relativamente aos seus parâmetros, a sua identificação dos respectivos parâmetros pode ser efectuada recorrendo a algoritmos de aprendizagem lineares.

De seguida são apresentadas algumas metodologias genéricas de identificação do modelo do tipo caixa negra ou cinzenta nos modelos híbridos série e paralelo.

### 5.3.1 Identificação do Modelo Híbrido Série

O processo de identificação de um modelo híbrido série necessita, para além dos conhecimentos dos primeiros princípios do sistema necessários para a definição do modelo fenomenológico, da existência de alguns conhecimentos fenomenológicos ou experimentais sobre a parte menos conhecida do sistema. A forma de interligação do modelo desconhecido com o modelo fenomenológico, a escolha dos regressores, o tipo de modelo (rede neuronal, etc.) e a sua estrutura interna dependem desses conhecimentos.

Oliveira em [Oli04] propõe duas estratégias de identificação de modelos híbridos série, qualquer que seja o tipo de modelo cinzento ou negro escolhido para a modelização da parte desconhecida do sistema:

- **Estratégia de identificação baseada no erro entre  $z$  e  $\hat{z}$**  (figura 5.4). Esta estratégia minimiza o erro quadrático entre a saída do modelo fenomenológico “inverso” e a saída do modelo do tipo caixa cinzenta ou negra.
- **Estratégia de identificação baseada no erro entre  $y$  e  $\hat{y}$**  (figura 5.5). Esta estratégia minimiza o erro entre a saída real do sistema e a saída do modelo híbrido série.

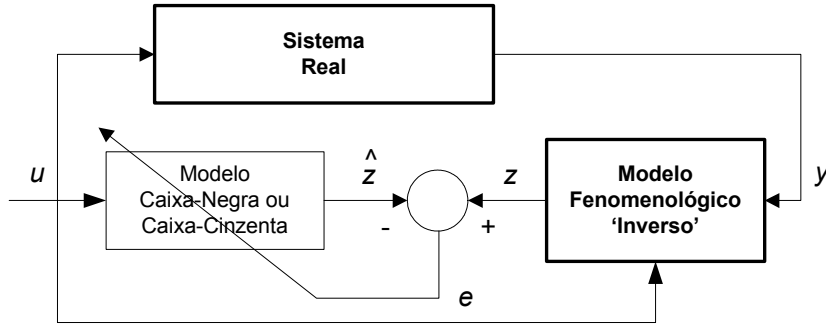


Figura 5.4: Identificação do modelo híbrido série baseada no erro entre  $\mathbf{z}$  e  $\hat{\mathbf{z}}$ .

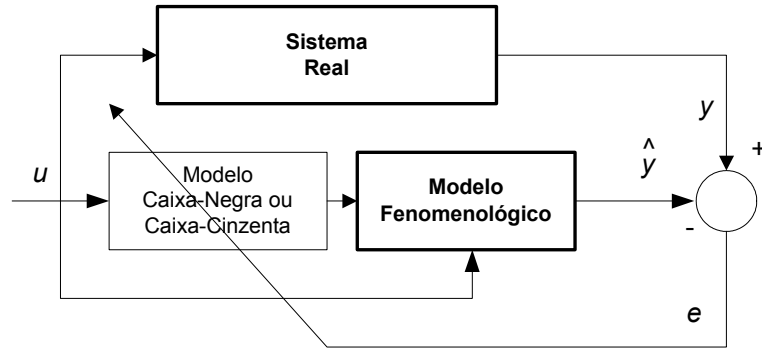


Figura 5.5: Identificação do modelo híbrido série baseada no erro entre  $\mathbf{y}$  e  $\hat{\mathbf{y}}$ .

### Estratégia de identificação baseada no erro entre $\mathbf{z}$ e $\hat{\mathbf{z}}$

Esta estratégia necessita de um conjunto de dados de treino do sistema  $(\mathbf{u}, \mathbf{z})$  para a identificação do modelo do tipo caixa cinzenta ou negra (figura 5.4). O problema está em obter os valores da saída  $\mathbf{z}$  do modelo desconhecido a partir do modelo fenomenológico.

Para o modelo híbrido série, definido na equação 5.3 como o vector da saída  $\hat{\mathbf{z}}$  do modelo desconhecido é tipicamente não mensurável, é possível calcular esse vector rescrevendo a equação 5.3 em ordem a  $\mathbf{z}$  (“invertendo” o modelo fenomenológico), como definido na equação 5.5:

$$\begin{aligned}\mathbf{z} &= f_{CCN}(\mathbf{x}, \mathbf{u}) = f_{PP}^{-1}(\dot{\mathbf{x}}, \mathbf{x}, \mathbf{u}) \\ \hat{\mathbf{y}} &= g(\mathbf{x})\end{aligned}\tag{5.5}$$

em que  $f_{PP}^{-1}(\cdot)$  é a função inversa de  $f_{PP}(\cdot)$  relativamente a  $\mathbf{z} = f_{CCN}(\cdot)$ .

Deste modo o cálculo do vector  $\mathbf{z}$  necessita do valor da variação de  $\mathbf{x}$  em ordem ao tempo. Uma aproximação possível para o cálculo da derivada de  $\mathbf{x}$  pode ser obtida utilizando um algoritmo numérico de diferenciação, como referido em [Oli04], permitindo deste modo obter o vector  $\mathbf{z}$  que conjuntamente com o vector das entrada  $\mathbf{u}$  permite a identificação do modelo do tipo caixa cinzenta ou negra  $f_{CCN}(\mathbf{x}, \mathbf{u})$ .



Outra alternativa para o cálculo do vector  $\mathbf{z}$  passa por transformar a equação 5.3 (ou 5.5) para o domínio discreto. Rescrevendo a função resultante no domínio discreto em ordem a  $\mathbf{z}$  e com os valores de saída real  $\mathbf{y}$  obtém-se o vector  $\mathbf{z}$  necessário para a identificação do respectivo modelo  $f_{CCN}(\mathbf{x}, \mathbf{u})$ .

Depois de obter os valores de  $\mathbf{z}$ , o modelo do tipo caixa cinzenta ou negra pode ser identificado minimizando, por exemplo, a função de custo  $E$ , definida na equação 5.6, baseada no erro quadrático médio (MSE) entre as saídas  $\mathbf{z}$  e  $\hat{\mathbf{z}}$ :

$$E|_{\theta} = \left\{ \frac{1}{M} \sum_{t=1}^M [z(k) - \hat{z}(k)]^T [z(k) - \hat{z}(k)] \right\} \quad (5.6)$$

em que  $z(k) - \hat{z}(k)$  é o erro entre a saída do modelo fenomenológico inverso e a saída do modelo do tipo caixa cinzenta ou negra,  $M$  é o número de elementos do vector dos dados de entrada e saída e  $\theta$  o vector dos parâmetros a identificar.

### Estratégia de identificação baseada no erro entre $\mathbf{y}$ e $\hat{\mathbf{y}}$

Esta estratégia necessita de um conjunto de dados de entrada e saída de treino do sistema  $(\mathbf{u}, \mathbf{y})$  (figura 5.5) que permitam a identificação dos parâmetros internos do modelo do tipo caixa negra ou cinzenta.

Pelo modelo fenomenológico definido na equação 5.3, verifica-se que para obter a saída  $\hat{\mathbf{y}}$  é necessário recorrer a um algoritmo de integração de  $\dot{\mathbf{x}}$ . Ao contrário da estratégia de identificação baseada no erro entre as saídas  $\mathbf{z}$  e  $\hat{\mathbf{z}}$ , que requer a utilização de um algoritmo de diferenciação numérica, a identificação baseada no erro entre as saídas  $\mathbf{y}$  e  $\hat{\mathbf{y}}$ , requer um algoritmo de integração numérica dos valores de  $\dot{\mathbf{x}}$ , como referido em [Oli04]. Deste modo, obtém-se uma aproximação dos valores da saída  $\hat{\mathbf{y}}$ .

A identificação do modelo do tipo caixa cinzenta ou negra é efectuada com base no conjunto de dados de entrada e saída de treino  $(\mathbf{u}, \mathbf{y})$ , minimizando a função  $E$  definida na equação 5.7:

$$E|_{\theta} = \left\{ \frac{1}{M} \sum_{k=1}^M [y(k) - \hat{y}(k)]^T [y(k) - \hat{y}(k)] \right\} \quad (5.7)$$

em que  $M$  é o número de elementos do vector dos dados de entrada e saída e  $\theta$  o vector dos parâmetros a identificar, baseada no erro quadrático médio entre as saídas  $\mathbf{y}$  e  $\hat{\mathbf{y}}$ .

Para a minimização da função de custo  $E$  (equação 5.7) podem ser utilizadas várias estratégias. Por exemplo, a optimização estocástica ou a minimização de  $E$  baseada no gradiente [TK94a]. O processo de optimização pode ser computacionalmente pesado, especialmente quando os modelos do tipo caixa cinzenta ou negra utilizados são complexos ou quando o conjunto de dados disponível é grande.

Psichiogios e Ungar [PU92b] sugerem a utilização do método da sensibilidade para a avaliação do gradiente  $\frac{\partial E}{\partial \theta}$  num modelo híbrido (balanço de massa combinado com uma rede neuronal artificial). O método da sensibilidade utiliza a informação do gradiente proveniente das equações de sensibilidade para a minimização da função de custo  $E$  identificando os parâmetros do modelo desconhecido.

Para o caso do modelo híbrido série genérico (equação 5.3) as equações de sensibilidade são obtidas desde que se garanta a continuidade e diferenciabilidade das funções  $f_{CCN}(\mathbf{x}, \mathbf{u}, \boldsymbol{\theta})$  utilizadas. A diferenciação de  $E$  em ordem aos parâmetros  $\boldsymbol{\theta}$  está definida na equação 5.8:

$$\frac{\partial E}{\partial \boldsymbol{\theta}} = \sum_{k=1}^M \left( \frac{\partial E}{\partial \hat{\mathbf{y}}} \right)_k \left( \frac{\partial \hat{\mathbf{y}}}{\partial \boldsymbol{\theta}} \right)_k = -\frac{2}{M} \sum_{k=1}^M [y(k) - \hat{y}(k)]^T \left( \frac{\partial \hat{\mathbf{y}}}{\partial \boldsymbol{\theta}} \right)_t \quad (5.8)$$

onde a matriz  $\left( \frac{\partial \hat{\mathbf{y}}}{\partial \boldsymbol{\theta}} \right)_k$  pode ser calculada através das equações de sensibilidade que surgem da diferenciação da equação 5.3 em ordem a  $\boldsymbol{\theta}$  [Oli04].

Depois de calculados os valores da matriz  $\left( \frac{\partial \hat{\mathbf{y}}}{\partial \boldsymbol{\theta}} \right)_k$  e utilizando o vector dos gradientes  $\frac{\partial E}{\partial \boldsymbol{\theta}}$ , é possível identificar o vector dos parâmetros  $\boldsymbol{\theta}$  de qualquer tipo de modelo cinzento ou negro (rede neuronal artificial ou sistema neuro-difuso) minimizando a equação 5.7.

Abonyi [AMS02] apresenta a identificação e o controlo de um CSTR (*continuous stirred tank reactor*) através de um modelo híbrido série identificado com base no erro entre as saídas  $\mathbf{y}$  e  $\hat{\mathbf{y}}$  e nas equações de sensibilidade. São utilizadas as equações no espaço de estados para a descrição física do processo e uma rede neuronal para modelizar a parte desconhecida do processo.

### 5.3.2 Identificação do Modelo Híbrido Paralelo

A identificação do modelo híbrido paralelo não exige conhecimentos sobre a parte desconhecida do sistema. Exige, no entanto, um conjunto de dados de entrada e saída representativo do sistema. Só assim é possível identificar correctamente o modelo do tipo caixa cinzenta ou negra, permitindo obter modelos híbridos paralelos com elevada capacidade de generalização.

O esquema de identificação dos parâmetros internos do modelo desconhecido do modelo híbrido paralelo está ilustrado na figura 5.6:

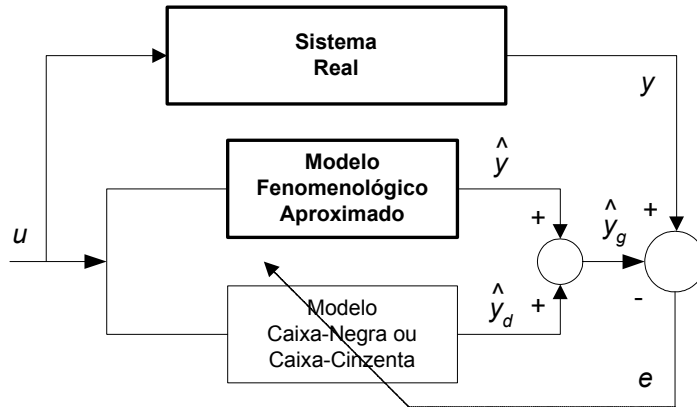


Figura 5.6: Identificação do modelo híbrido paralelo.

Os dados de entrada e saída necessários para a identificação do modelo do tipo

caixa cinzenta ou negra são  $(\{\mathbf{u}, \mathbf{y} - \hat{\mathbf{y}}\})$ , em que  $\mathbf{u}$  é o vector das entradas,  $\mathbf{y}$  é o vector das saídas reais e  $\hat{\mathbf{y}}$  é o vector das saídas estimadas pelo modelo fenomenológico aproximado.

Através do modelo fenomenológico aproximado genérico definido na equação 5.2 verifica-se que, para obter o vector da saída  $\hat{\mathbf{y}}$ , é necessário recorrer a um algoritmo de integração de  $\dot{\mathbf{x}}$ , como no caso do modelo híbrido série. Deste modo, obtém-se uma aproximação dos valores da saída  $\hat{\mathbf{y}}$  necessários para a identificação do modelo desconhecido.

Outra alternativa para obter o vector das saídas estimadas  $\hat{\mathbf{y}}$  do modelo fenomenológico passa por transformar a equação 5.2 para o domínio discreto permitindo o cálculo do vector das saídas  $\hat{\mathbf{y}}$ .

A identificação do modelo do tipo caixa cinzenta ou negra é efectuada com base no conjunto de dados de entrada e saída do sistema  $(\{\mathbf{u}, \mathbf{y} - \hat{\mathbf{y}}\})$ , minimizando a função de custo  $E$ , definida na equação 5.9, baseada no erro quadrático médio entre a saída do sistema  $\mathbf{y}$  e a saída estimada do modelo híbrido paralelo  $\hat{\mathbf{y}}_g = \hat{\mathbf{y}} - \hat{\mathbf{y}}_d$ :

$$E|_{\theta} = \left\{ \frac{1}{M} \sum_{k=1}^M (y(k) - \hat{y}_g(k))^T (y(k) - \hat{y}_g(k)) \right\} \quad (5.9)$$

em que  $M$  é o número de elementos do vector dos dados de entrada e saída e  $\theta$  é o vector dos parâmetros a identificar.

Hussain [NH04] modeliza e identifica um processo de polimerização utilizando a modelização híbrida série e paralela, em que o modelo desconhecido é modelizado por uma rede neuronal. No caso da modelização híbrida paralela, a saída do modelo fenomenológico é aproximada aplicando um método de linearização e integração numérica, obtendo um modelo aproximado. Com esta saída e com as entradas e saídas do sistema real, a rede neuronal artificial é identificada.

Outro caso da aplicação da modelização híbrida paralela na modelização de um reactor exotérmico é apresentado em [XJ02]. A diferença entre o modelo fenomenológico aproximado discreto e a saída real do sistema é modelizada utilizando uma rede neuronal colocada em paralelo com o modelo fenomenológico.

## 5.4 Conclusão

Este capítulo apresentou dois métodos de modelização construídos essencialmente com base em conhecimentos fenomenológicos do sistema: a modelização híbrida série e híbrida paralela. Os modelos híbridos são construídos com base nas equações de balanço de um sistema cujas relações fenomenológicas nem sempre são completamente conhecidas. Para ultrapassar este problema, os modelos híbridos combinam em série, ou em paralelo, com o modelo fenomenológico, um modelo do tipo caixa cinzenta ou negra que irá modelizar as partes mal conhecidas ou desconhecidas.

Este tipo de modelos maximiza a incorporação de conhecimentos *a priori* sobre o sistema, permitindo a modelização de sistemas não lineares baseada na modelização fenomenológica conjugada com a modelização cinzenta ou negra. Permite ainda a

construção de modelos de complexidade matemática relativamente baixa, com elevado nível de interpretabilidade e com elevada capacidade de generalização.

Na secção 5.2 apresentou-se uma descrição dos modelos híbridos série e paralelo. Apresentou-se um método genérico para a definição das equações fenomenológicas baseadas nas equações de balanço do sistema. Definiram-se os modelos híbridos série e paralelo e apontaram-se dois métodos para a modelização da função desconhecida do modelo fenomenológico: redes neuronais e sistemas neuro-difusos.

As técnicas de identificação dos modelos híbridos série e paralelo foram apresentadas e descritas na secção 5.3. Para a identificação do modelo híbrido série foram apresentadas duas estratégias. A primeira é baseada na minimização do erro entre a saída do modelo do tipo caixa cinzenta ou negra e do valor calculado através do modelo fenomenológico “inverso” e a segunda, é baseada na minimização do erro entre as saídas do modelo híbrido série e do sistema real. Para a identificação do modelo híbrido paralelo foi apresentada a estratégia baseada na minimização da diferença entre as saídas do sistema real e do modelo híbrido paralelo.

# Capítulo 6

## Estruturas de Controlo Baseadas em Modelos

### 6.1 Introdução

Após terem sido apresentados e discutidos os três métodos de modelização e identificação cinzenta abordadas neste trabalho, este capítulo explora várias formas de utilizar estes modelos para o controlo automático de sistemas.

As estruturas controlo baseadas em modelos abordadas neste capítulo são: Controlo com Modelo Inverso, Controlo Aditivo, Controlo Baseado no Modelo Interno, Controlo Smith Preditivo e Controlo Preditivo Generalizado. As três últimas estruturas são, actualmente, as mais utilizadas no controlo de sistemas não lineares [Abo03], [Sou03], [VM04f], [VM04b], [VDM04a] e [PA88].

As malhas de controlo apresentadas neste capítulo permitem o controlo de sistemas lineares e não lineares desde que se garanta a existência dos modelos directos e inversos e a sua estabilidade.

### 6.2 Controlo com Modelo Inverso

O controlador com modelo inverso (*Direct Inverse Control - DIC*) possui uma estrutura simples. Consiste apenas na ligação em série do modelo inverso com o sistema a controlar, como ilustra a figura 6.1:

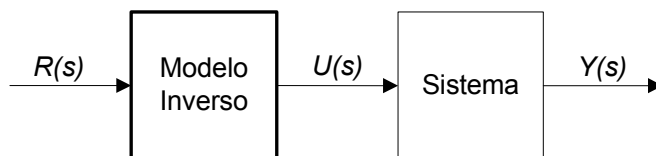


Figura 6.1: Controlador com modelo inverso.

Considerando  $S$  a função transferência do sistema,  $M$  e  $M^{-1}$ , respectivamente, as funções transferências dos modelos directo e inverso no domínio de Laplace, se o modelo

inverso for de “boa qualidade” (erro de modelização muito baixo), a saída do sistema  $Y(s)$  será muito semelhante à entrada  $R(s)$ , uma vez que a malha de controlo tem como função transferência  $M^{-1}S = 1$ , em que  $M \cong S$ .

O controlador com modelo inverso representado no domínio discreto está ilustrado no diagrama de blocos da figura 6.2, em que as variáveis  $r(\cdot)$ ,  $u(\cdot)$  e  $y(\cdot)$  estão definidas por  $r(k+1)$ ,  $u(k)$  e  $y(k+1)$ , respectivamente, por forma a representar a existência de um tempo morto de, pelo menos, uma amostra entre a entrada e a saída do sistema. Este procedimento é utilizado nos vários diagramas de blocos de modelos e controladores apresentados ao longo desta dissertação.

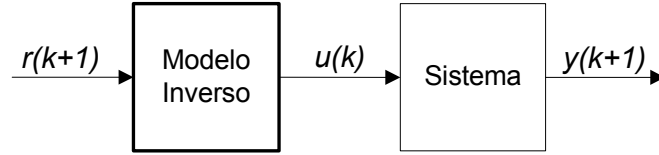


Figura 6.2: Controlador com modelo inverso.

Designando por  $m(\cdot)$  e  $m^{-1}(\cdot)$  os modelos discretos directo e inverso relativamente aos modelos  $M$  e  $M^{-1}$ , a estimação da saída do modelo directo  $\hat{y}(k+1)$  é função de  $m(\cdot)$  e de um conjunto de regressores, como definido na equação 6.1:

$$\hat{y}(k+1) = m(y(k), \dots, y(k - n_y + 1), u(k), \dots, u(k - n_u + 1)) \quad (6.1)$$

em que  $n_y$  é o número de regressores da saída  $y$  e  $n_u$  é o número de regressores da entrada  $u$ . A estimação da saída do modelo inverso  $\hat{u}$  é função de  $m^{-1}(\cdot)$  e de um conjunto de regressores, como definido na equação 6.2:

$$\hat{u}(k) = m^{-1}(y(k+1), y(k), \dots, y(k - n_y + 1), u(k-1), \dots, u(k - n_u + 1)) \quad (6.2)$$

Em sistemas de 1ª ordem, as saídas dos modelos directo e inverso são estimadas com base em dois regressores como definido, respectivamente, nas equações 6.3 e 6.4:

$$\hat{y}(k+1) = m(y(k), u(k)) \quad (6.3)$$

$$\hat{u}(k) = m^{-1}(y(k+1), y(k)) \quad (6.4)$$

Para sistemas de 2ª ordem, as saídas dos modelos directo e inverso são estimadas com base em quatro regressores como definido, respectivamente, nas equações 6.5 e 6.6:

$$\hat{y}(k+1) = m(y(k), y(k-1), u(k), u(k-1)) \quad (6.5)$$

$$\hat{u}(k) = m^{-1}(y(k+1), y(k), y(k-1), u(k-1)) \quad (6.6)$$

Os diagramas de blocos das estruturas de controlo com modelo inverso para sistemas de 1ª ordem e de 2ª ordem estão apresentados, respectivamente, nas figuras 6.3 e 6.4,

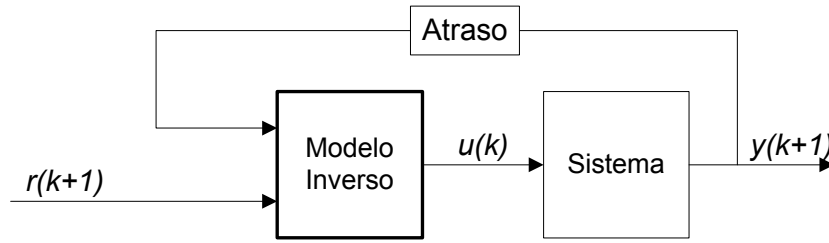


Figura 6.3: Controlador com modelo inverso para sistemas de 1ª ordem.

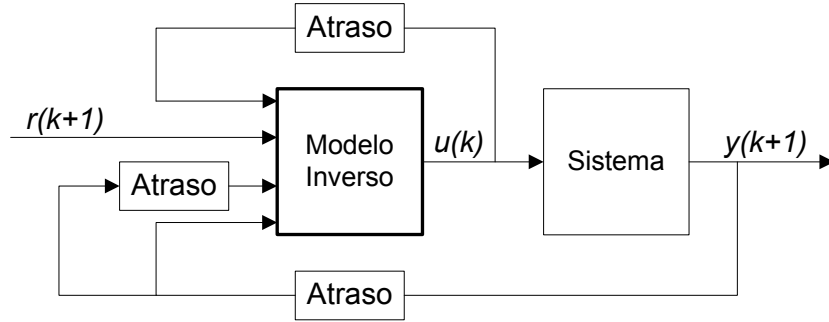


Figura 6.4: Controlador com modelo inverso para sistemas de 2ª ordem.

em que o sinal  $r(k+1)$  é a referência aplicada ao sistema em malha fechada e *Atraso* representa o atraso de uma amostra.

Nas equações 6.4 e 6.6,  $y(k+1)$  não está disponível na iteração  $k$ , é habitual ser substituído pela referência correspondente  $r(k+1)$ . Como se pode ver pelas figuras 6.3 e 6.4, estas estruturas não possuem malhas de realimentação do erro de controlo. A única forma de realimentação existente é efectuada através dos regressores aplicados ao modelo inverso.

Esta estrutura de controlo, apesar de muito simples, pode não atingir resultados de controlo satisfatórios. Se o modelo inverso não for de boa qualidade  $M^{-1}S \neq 1$ , os resultados de controlo também não o serão. Neste situação a malha de controlo com modelo inverso vai apresentar erro em estado estacionário.

## 6.3 Controlo Aditivo

O controlador aditivo (*Additive Feedforward Control - AFC*) consiste, como mostra a figura 6.5, na adição de um modelo inverso colocado em paralelo com um controlador já existente.

Esta malha de controlo é, normalmente, utilizada para melhorar o desempenho de um controlador já existente cujos resultados de controlo não são completamente satisfatórios.

Tome-se, por exemplo, um sistema linear variável com o ponto de funcionamento de forma conhecida a ser controlado por um controlador PID (*Proportional Integral*

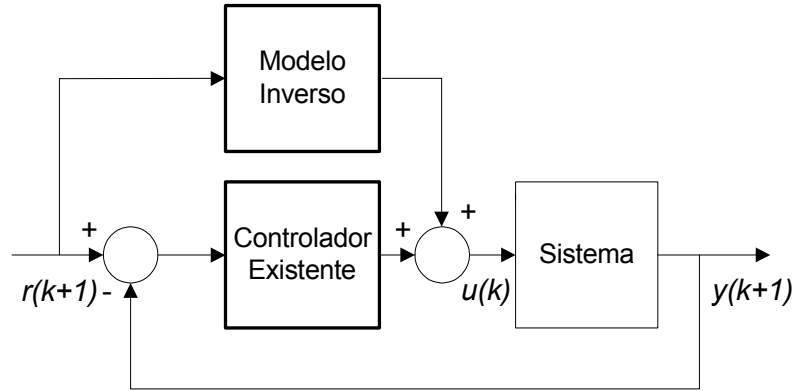


Figura 6.5: Controlador aditivo.

*Derivative*) de parâmetros fixos (controlador existente). Neste caso o desempenho do controlador não é sempre igual e satisfatório em todos os pontos de funcionamento. A introdução do modelo inverso do sistema em paralelo vai melhorar o desempenho deste controlador.

Os problemas apontados à malha de controle com modelo inverso são até certo ponto eliminados com a utilização desta malha, uma vez que, para além do modelo inverso esta malha possui um segundo controlador colocado em paralelo com o modelo inverso que compensa os desvios provocados pela imprecisão do modelo inverso ou por cargas que actuem no sistema.

Para o caso de um sistema de 2ª ordem, o diagrama de blocos do controlador aditivo puro com as respectivas ligações dos regressores está ilustrado na figura 6.6. O diagrama de blocos deste controlador para sistemas de 1ª ordem é semelhante ao ilustrado na figura 6.6, eliminando os dois regressores  $y(k-1)$  e  $u(k-1)$  aplicados no bloco do modelo inverso.

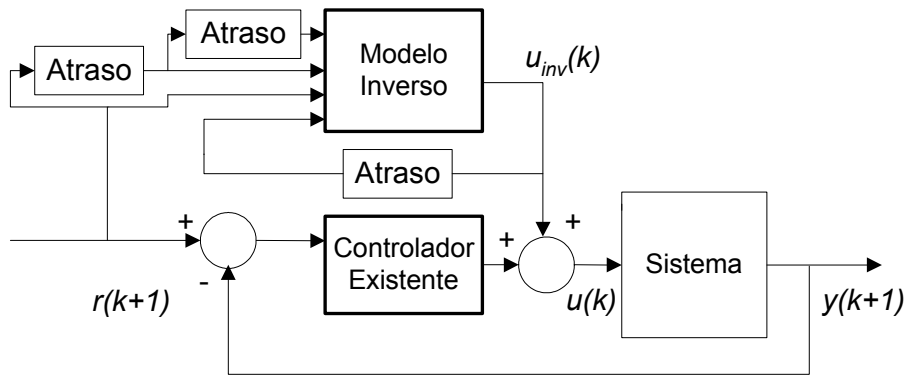


Figura 6.6: Controlador aditivo puro para sistemas de 2ª ordem.

A estrutura de controle apresentada no diagrama de blocos da figura 6.6 é denominada por controlador aditivo puro [Sor94]. Nesta solução, as realimentações do modelo



inverso são versões atrasadas do sinal de referência e do sinal de saída do modelo inverso. A utilização de um modelo inverso de baixa qualidade implica a redução da eficácia e robustez desta malha de controlo.

Existe ainda outra solução, ilustrada na figura 6.7, denominada controlador aditivo misto [Sor94], que consiste na utilização de versões atrasadas do sinal de saída real e do sinal de controlo para a realimentação do modelo inverso. Esta solução permite a construção de uma malha de controlo mais eficiente e robusta.

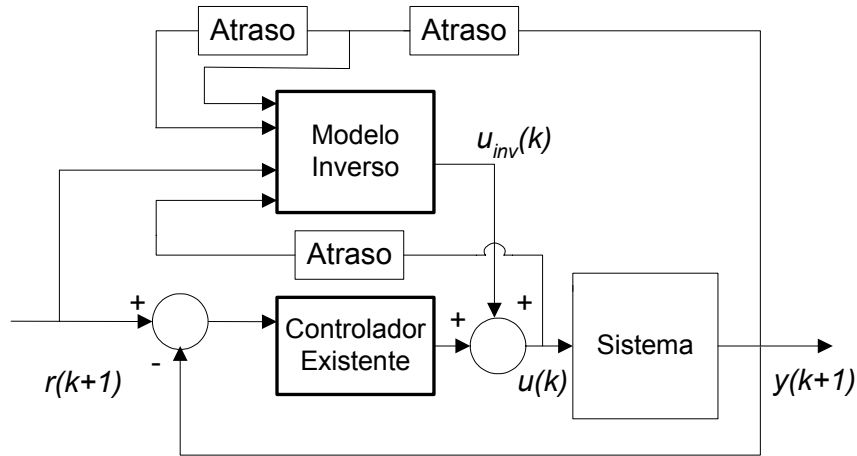


Figura 6.7: Controlador aditivo misto para sistemas de 2ª ordem.

## 6.4 Controlo Baseado no Modelo Interno

O controlador baseado no modelo interno (*Internal Model Control - IMC*) [GM82], [EMP86] é constituído, como ilustra a figura 6.8, por um controlador genérico e por um modelo directo colocado em paralelo com o sistema. A diferença entre as respectivas saídas  $e_r(k+1)$  é realimentada negativamente para a entrada do controlador. Esta estrutura de controlo é uma das mais populares e mais utilizadas no controlo baseado em modelo.

A malha de controlo baseada no modelo interno pode utilizar como controlador o respectivo modelo inverso. Neste caso, para um bom desempenho de controlo, é necessário que o modelo directo esteja emparelhado com o modelo inverso do sistema, ou seja  $M^{-1}M = 1$ .

É normal utilizar um filtro passa-baixo na malha de realimentação aplicado ao sinal  $e_r(k+1)$  ou ao sinal de referência  $r(k+1)$  ou ainda aplicado a ambos os sinais referidos, como apresentado na figura 6.8. Este filtro tem como objectivo “suavizar” o efeito de perturbações, erros do modelo e transições bruscas do sinal de referência.

O filtro passa-baixo utilizado tem como equação:

$$w(k+1) = \alpha(r(k+1) - e_r(k+1)) + (1 - \alpha)w(k) \quad (6.7)$$

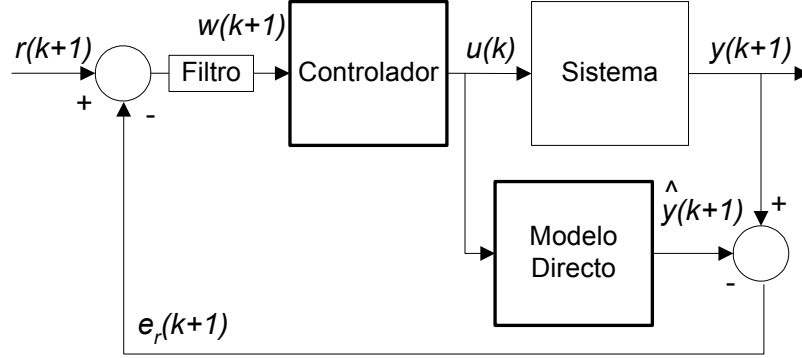


Figura 6.8: Controlador com modelo interno.

onde  $w(k+1)$  é o sinal filtrado a aplicar no modelo inverso e  $\alpha$  define a constante de tempo do filtro com  $0 \leq \alpha < 1$ . Este parâmetro deve ser escolhido de modo a que malha de controlo respeite um compromisso entre desempenho e robustez desejadas [GM82].

Para o caso de um sistema de 2ª ordem, o diagrama de blocos do controlador baseado no modelo interno, cujo controlador é o respectivo modelo inverso, está ilustrado na figura 6.9. O diagrama de blocos deste controlador para sistemas de 1ª ordem é semelhante ao ilustrado na figura 6.9, eliminando os dois regressores  $\hat{y}(k-1)$  e  $u(k-1)$  aplicados no bloco do modelo inverso:

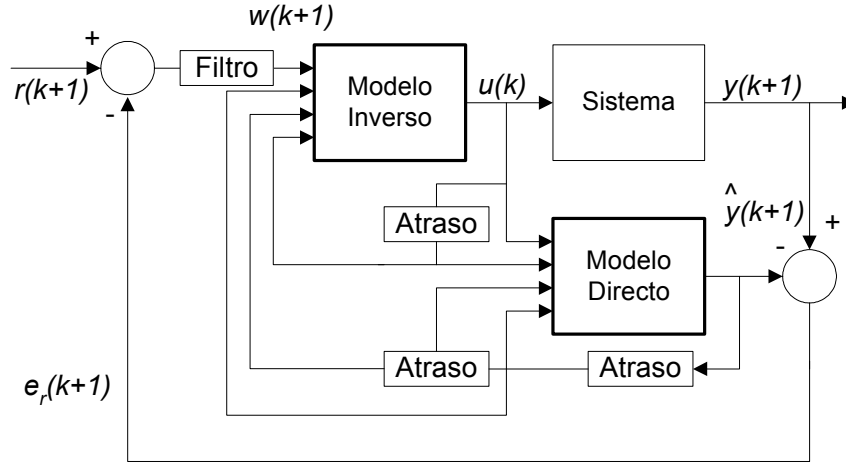


Figura 6.9: Controlador com modelo interno para modelos de 2ª ordem.

#### 6.4.1 Controlo Smith Preditivo Baseado no Modelo

Um caso particular do controlador baseado no modelo interno é o controlador Smith preditivo (*Smith Predictive Control - SPC*) [PA88]. O diagrama de blocos deste controlador está ilustrado na figura 6.10, em que  $d$  é o tempo morto do sistema.

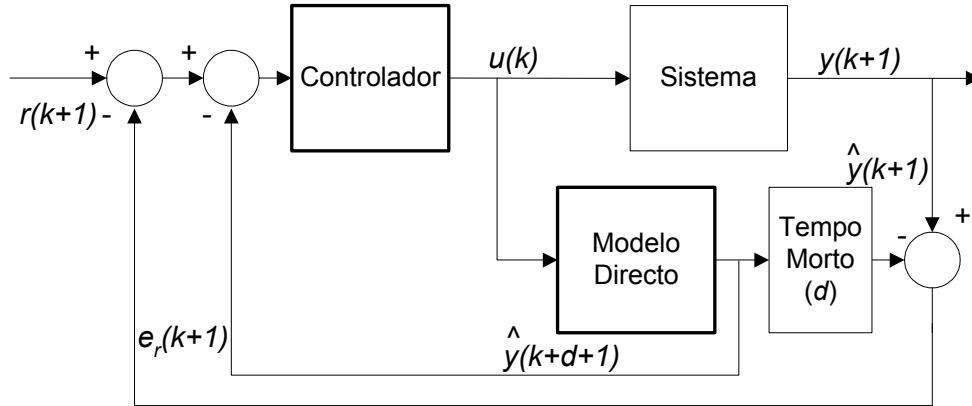


Figura 6.10: Controlador Smith preditivo baseado no modelo interno.

Neste caso, o modelo directo do sistema não inclui o tempo morto, sendo este modelizado à parte. Esta característica permite ao controlador prever antecipadamente ( $d$  iterações antes) o comportamento do sistema. Com esta informação o controlador vai poder actuar antecipadamente de modo a minimizar o erro da malha de controlo.

Os controladores Smith preditivos são adequados para o controlo de sistemas com tempo morto elevado [PA88]. O desempenho desta malha de controlo não é afectado pelo valor do tempo morto, desde que este seja um múltiplo inteiro do período de amostragem. Quando o tempo morto possui uma parte fraccionária o desempenho do controlador degrada-se. Este controlador exhibe sobre-elevação quando o tempo morto real é maior do que o tempo morto  $d$  e sub-elevação quando o tempo morto real é menor do que o tempo morto  $d$  [TC99].

Para o caso de um sistema de 2ª ordem, o diagrama de blocos do controlador Smith preditivo com as respectivas ligações dos regressores está ilustrado na figura 6.11. O diagrama de blocos deste controlador para sistemas de 1ª ordem é semelhante ao ilustrado na figura 6.11, eliminando os dois regressores  $\hat{y}(k+d-1)$  e  $u(k-1)$  aplicados no bloco com o modelo inverso:

## 6.5 Controlo Preditivo Generalizado

O controlador preditivo generalizado ou preditivo baseado em modelo (*Generalized Predictive Control - GPC or Model Predictive Control - MPC*) [CMT87], [CB99] é uma estrutura de controlo que utiliza a predição do sinal de saída do sistema para calcular o sinal de controlo.

A ideia básica é calcular uma sequência de amostras futuras dos sinais de controlo de modo a minimizar uma função de custo definida num horizonte de controlo. O cálculo das amostras futuras dos sinais de controlo, num determinado horizonte, é baseado no modelo directo do sistema e na sequência futura do sinal de referência. O princípio básico do controlo preditivo está representado na figura 6.12:

As saídas futuras para um determinado horizonte de predição  $h_p$  são previstas a cada amostra  $k$  usando o modelo do sistema. Os valores de saída preditos  $\hat{y}(k+j)$ ,  $j = 1..h_p$

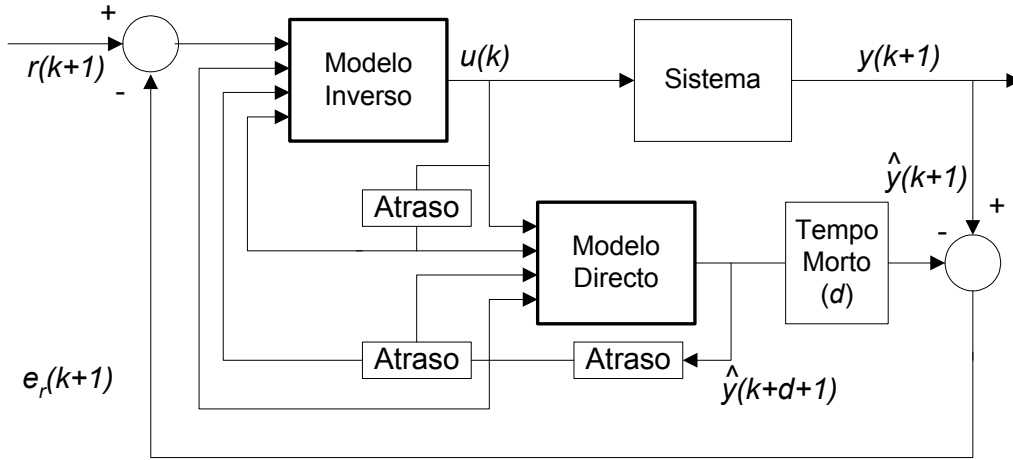


Figura 6.11: Controlador Smith preditivo para modelos de 2ª ordem.

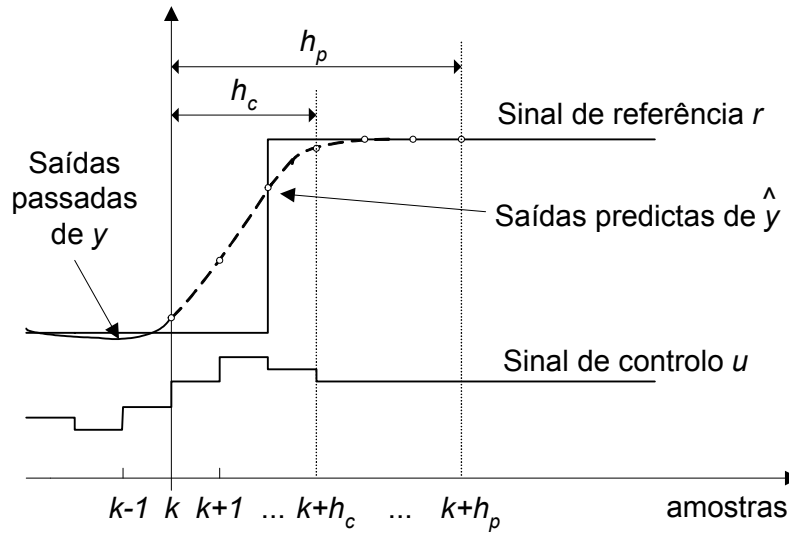


Figura 6.12: Princípio básico do controlo preditivo.

dependem do ponto de funcionamento do sistema na amostra  $k$  e dos sinais de controlo futuros  $u(k+j)$ ,  $j = 1..h_c$ , onde  $h_c$  é o horizonte de controlo. O sinal de controlo só varia no horizonte de controlo, permanecendo constante no restante horizonte de predição  $u(k+j) = u(k+h_c-1)$ ,  $j = h_c..h_p-1$ .

O horizonte de controlo está relacionado com a ordem do modelo e é sempre menor do que o horizonte de predição. O horizonte de predição está normalmente relacionado com o tempo de resposta do sistema. Para sistemas não lineares o tempo de resposta pode variar implicando a estimação de um valor médio.

A sequência de sinais de controlo futuros é obtida pela optimização de uma função de custo  $J(h_p, h_c)$  que descreve o objectivo de controlo. No GPC clássico a função de custo é normalmente dada pelas equações 6.8 ou 6.9:

$$J(h_p) = \sum_{i=1}^{h_p} \left( \alpha(i) (r(k+i) - \hat{y}(k+i))^2 + \beta(i) (\Delta u(k+i-1))^2 \right) \quad (6.8)$$

$$J(h_p, h_c) = \sum_{i=1}^{h_p} \alpha(i) (r(k+i) - \hat{y}(k+i))^2 + \sum_{i=1}^{h_c} \beta(i) (\Delta u(k+i-1))^2 \quad (6.9)$$

em que  $r(k+i) - \hat{y}(k+i)$  com  $i = 1..h_p$  são os erros futuros do controlador,  $\Delta u(k+i)$  são os incrementos futuros do sinal de controlo e  $\alpha(i)$  e  $\beta(i)$  são, respectivamente, os pesos atribuídos ao erro e ao incremento do sinal de controlo (normalmente valores constantes).

Para sistemas com tempo morto, como mostra a figura 6.13, o horizonte de controlo não é de  $h_p$ , mas sim de  $h_{p1}$  até  $h_{p2}$ , em que  $h_{p1}$  é o horizonte de predição mínimo directamente relacionado com o tempo morto do sistema (podendo ser igual) e  $h_{p2}$  é o horizonte de predição máximo:

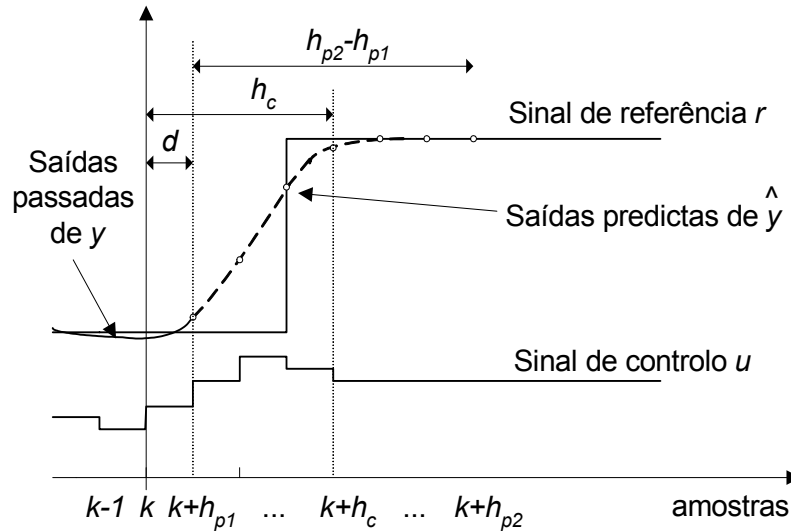


Figura 6.13: Princípio básico do controlo preditivo com tempo morto.

Neste caso, a função de custo  $J(h_{p1}, h_{p2}, h_c)$  a otimizar passa a ser definida pela equação 6.10:

$$J(h_{p1}, h_{p2}, h_c) = \sum_{i=h_{p1}}^{h_{p2}} \alpha(i) (r(k+i) - \hat{y}(k+i))^2 + \sum_{i=1}^{h_c} \beta(i) (\Delta u(k+i-1))^2 \quad (6.10)$$

O diagrama de blocos do controlador preditivo generalizado está ilustrado na figura 6.14.

Este controlador determina o próximo sinal de controlo a aplicar no sistema optimizando uma das funções de custo atrás apresentadas com base no respectivo modelo directo do sistema e no sinal gerado pelo bloco de referência  $w(k+1)$ . O bloco gerador de referência tem como função filtrar o sinal  $r(k+1) - y(k+1)$  definindo o novo sinal de

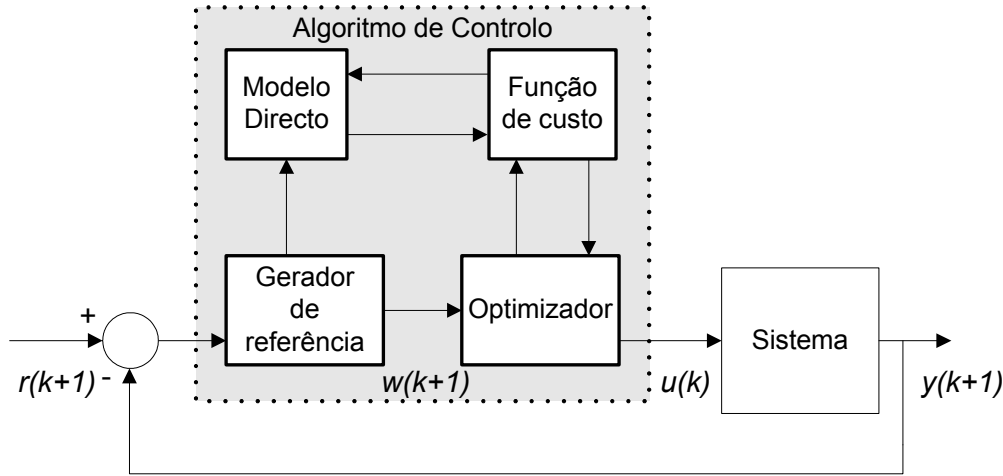


Figura 6.14: Controlador preditivo generalizado.

referência  $w(k+1)$ . Normalmente,  $w(k+i)$  substitui a referência  $r(k+i)$  com  $i = 1..h_{p2}$  na equação 6.9 ou  $i = h_{p1}..h_{p2}$  na equação 6.10 [Sou03].

Esta malha de controlo apresenta diversas vantagens das quais se destacam as seguintes [Abo03]:

- Os tempos mortos elevados são tratados de uma forma explícita;
- A manipulação e restrição das variáveis de estado é tratada de uma forma explícita;
- Sistemas de fase não mínima são tratados facilmente (requer um horizonte de controlo mínimo,  $h_{p1} > 1$ );
- O conhecimento das mudanças futuras do sinal de referência é incluído sendo este muito útil no controlo pré-programado de sistemas;

Este controlador pode ser conjugado com o controlador baseado no modelo interno [GM82] ou com o controlador Smith preditivo, obtendo diferentes malhas de controlo com diferentes características.

O problema da aplicação do controlo preditivo generalizado em sistemas não lineares é a resolução, em cada período de amostragem, da optimização não linear da função de custo. A elevada complexidade matemática pode restringir a aplicação deste algoritmo de controlo quando executado por sistemas computacionais de baixo poder de cálculo.

A resolução do problema da optimização não linear pode ser resolvida através de vários métodos [Hen98]:

- **Sequencial.** O modelo não linear é aplicado directamente na optimização da função de custo. A optimização e as equações do modelo não linear são calculadas iterativamente até se atingir a exactidão desejada.

- **Simultâneo.** A execução do problema da optimização e a resolução das equações do modelo não linear são efectuadas em simultâneo. Esta técnica envolve a transformação do modelo dinâmico em equações algébricas usando técnicas residuais pesadas.
- **Linearização sucessiva.** Este é o método mais simples de tratar as equações do modelo, extraindo o modelo linear local à volta do ponto de funcionamento, deste modo, o problema fica reduzido à optimização MPC linear.
- **Linearização alternativa.** Para melhorar a aplicabilidade de MPC não linear são propostas alternativas de formulação do problema para melhorar as propriedades computacionais, como o método FHMP (Fuzzy Hammerstein Model-based Predictive Control - FHMP) proposto em [Abo03], que em primeiro lineariza o sistema através de um modelo inverso e segundo, utiliza o seu modelo linear para a optimização da função de custo utilizando técnicas lineares.

J. Sousa [Sou03] propõe um controlador preditivo baseado na inversão do modelo difuso directo em que a optimização da função de custo é efectuada utilizando dois algoritmos diferentes: o algoritmo iterativo e sequencial *Branch-and-Bound* - *B&B* e algoritmos genéticos. O autor concluiu que a utilização de algoritmos genéticos implica uma carga computacional muito maior, particularmente no caso de horizontes de controlo longos. Quanto à qualidade de controlo, os dois métodos obtêm resultados semelhantes.

Abonyi apresenta o FHMP [Abo03] que recorre ao algoritmo de linearização alternativa. Em primeiro lugar, o algoritmo lineariza o sistema através de um modelo inverso; em segundo, utiliza o modelo linearizado aproximado para a optimização da função de custo como no GPC linear [CB99].

## 6.6 Conclusão

Este capítulo apresentou de forma abreviada várias malhas de controlo baseadas em modelos. Actualmente as malhas mais utilizadas no controlo de sistemas lineares e/ou não lineares são: controlo baseado no modelo interno e controlo preditivo generalizado, uma vez que estas possuem características específicas como, a predição antecipada do comportamento dos sistemas, permitindo um controlo mais eficiente.

A escolha da estrutura de controlo a utilizar é condicionada pelo tipo de modelo utilizado e pelos requisitos de complexidade do algoritmo exigidos.





# Capítulo 7

## O Esquentador Doméstico

### 7.1 Introdução

Este capítulo apresenta, descreve e caracteriza um esquentador doméstico genérico. Este equipamento funciona a gás e fornece em tempo real um determinado caudal de água quente para uso doméstico.

O esquentador é um sistema relativamente complexo sendo definido por três variáveis de entrada e uma de saída (sistema MISO) e caracterizado por um comportamento fortemente não linear. Este termodoméstico vai servir como caso prático para o teste e avaliação dos vários métodos de modelização e controlo apresentados nos capítulos anteriores.

Este capítulo caracteriza o esquentador ao nível mecânico, funcional, estático e dinâmico, descrevendo ainda a plataforma de testes implementada. Esta caracterização divide-se na descrição da estrutura física, na definição das variáveis de entrada e saída e respectivas gamas de variação, na descrição dos vários estados de funcionamento e no estudo do comportamento estático e dinâmico. Por fim, este capítulo apresenta a plataforma de testes constituída por um esquentador, uma placa de controlo e um computador pessoal que será utilizada para testar e avaliar as várias soluções de controlo do esquentador.

### 7.2 Estrutura Física

O esquentador utilizado neste trabalho é constituído por vários elementos mecânicos, sensores e actuadores. A figura 7.1 mostra, de uma forma esquemática, os vários componentes no sistema.

Os elementos mecânicos fundamentais são: o queimador de gás, a câmara de combustão, o permutador e a caixa metálica de isolamento térmico. Estes elementos são os componentes mecânicos mínimos exigidos para o funcionamento de um esquentador.

Além dos elementos mecânicos já referidos, o esquentador possui um conjunto de sensores e actuadores. Os sensores existentes no esquentador são: duas resistências de coeficiente de temperatura negativo (NTC), um sensor de turbina, um pressostato, um termóstato e um bipolo de ionização. Estes sensores destinam-se a medir/verificar,

respectivamente, as seguintes grandezas/fenómenos: temperatura da água quente, temperatura da água fria, caudal da água, exaustão, excesso de temperatura e ionização. Os actuadores existentes no esquentador são: válvula de gás proporcional, válvula de gás *on-off*, faiscador e ventilador. Estes actuadores destinam-se a controlar, respectivamente, as seguintes grandezas/fenómenos: fluxo de gás que entra no esquentador, a combustão e a exaustão dos gases de combustão.

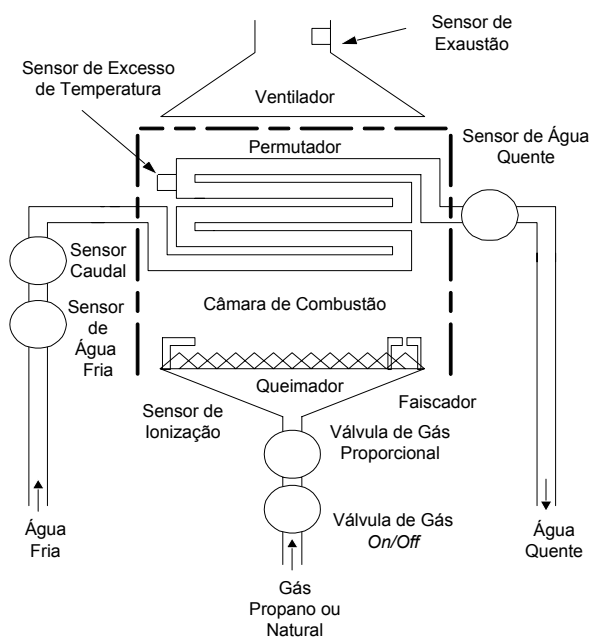


Figura 7.1: Componentes constituintes de um esquentador doméstico.

### 7.2.1 Elementos Mecânicos

Os elementos mecânicos existentes num esquentador são: o queimador de gás, a câmara de combustão, o permutador e a caixa metálica de isolamento térmico. O queimador de gás (figura 7.2) é constituído por linhas paralelas de injectores de gás que servem para distribuir e queimar o gás, transformando a energia fóssil em energia térmica. O queimador pode funcionar com gás propano, butano ou natural sendo para isso necessário substituir os respectivos injectores.

A câmara de combustão é de cobre (figura 7.3 - parte interior não visível da caixa de cobre que contém ar) e serve para alojar a combustão e para encaminhar o ar quente até ao permutador.

O permutador de calor, também de cobre, de formato laminado (figura 7.3 - parte superior) é responsável pela permutação da energia térmica do ar produzida pela combustão para a energia térmica da água que se encontra no seu interior.

O queimador de gás, a câmara de combustão e o permutador são colocados dentro de uma caixa metálica revestida por uma fibra isolante para evitar perdas de energia para o exterior. A minimização das perdas de energia para o exterior maximiza o rendimento do esquentador.



Figura 7.2: Queimador de gás



Figura 7.3: Câmara de combustão e permutador

### 7.2.2 Sensores

Os sensores existentes no esquentador são: duas resistências de coeficiente negativo (NTC), um sensor de turbina, um pressostato, um termóstato e um bipolo de ionização.

As duas resistências de coeficiente negativo são sensores de temperatura e destinam-se, respectivamente, a medir as temperaturas da água à saída e à entrada do esquentador.

O sensor de turbina é um dispositivo que permite medir o caudal de água que circula dentro do permutador de calor.

O pressostato é um sensor de pressão utilizado para verificar a existência de exaustão no esquentador. Este sensor, ao detectar nos seus terminais duas pressões diferentes (dois tubos colocados em sítios diferentes sentindo pressões de ar diferentes), estabelece um contacto físico entre dois pontos. O contacto físico, aberto ou fechado, destina-se a detectar se o esquentador está, ou não, a ser ventilado.

O termóstato é um sensor de temperatura utilizado para verificar a existência de excesso de temperatura no permutador. Este sensor, ao atingir uma dada temperatura, altera o contacto físico entre dois pontos, detectando assim, se a temperatura do permutador está acima de um determinado valor.

O bipolo de ionização, visível no canto inferior direito da figura 7.2, é um sensor resistivo sensível à existência de combustão entre os seus terminais. Esse sensor é constituído por um condutor eléctrico colocado a cerca de um centímetro da chama do queimador. A resistência entre o fio condutor e a carcaça do esquentador irá funcionar como uma resistência variável. Esta resistência ou possui um valor elevado ou é infinita correspondendo à existência, ou não, de chama no queimador. Com esta variação é possível detectar a existência de combustão.

### 7.2.3 Actuadores

Os actuadores existentes no esquentador são: válvula proporcional, válvula *on-off*, faiscador e ventilador.

As duas válvulas estão incorporadas num bloco SIT 845 SIGMA (ver figura 7.4). Este bloco é constituído por dois solenóides, sendo um controlado por corrente através de um sinal modulado por largura de impulso (PWM) e o outro de funcionamento *on/off*. A válvula proporcional serve para controlar a quantidade de gás fornecido ao queimador a cada momento. A válvula *on/off* serve para abrir ou fechar o gás fornecido ao queimador.



Figura 7.4: Válvulas de gás.

O faiscador, visível no canto inferior esquerdo da figura 7.2, destina-se a gerar uma faísca junto do queimador provocando o início da combustão. Este actuador é um conjunto de dois condutores separados por quatro milímetros, localizados junto ao queimador, nos quais se aplica uma tensão alternada e momentânea de 10 KVolt.

O ventilador é responsável pela exaustão dos gases de combustão. Este actuador é um motor de corrente alternada acoplado a uma ventoinha que impõe a exaustão dos gases de combustão a uma velocidade constante.

## 7.3 Definição das Entradas e Saídas

O esquentador é, como representado na figura 7.5, um sistema com várias entradas e uma única saída - sistema MISO (*Multiple Input Single Output*).

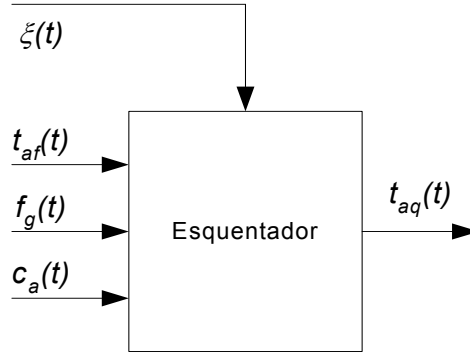


Figura 7.5: Variáveis de entrada e de saída do esquentador

As variáveis de entradas são: temperatura da água que entra no esquentador - temperatura da água fria  $t_{af}(t)$ , caudal de água  $c_a(t)$  que atravessa o esquentador e o fluxo de gás aplicado ao queimador do esquentador - fluxo de gás  $f_g(t)$ .

A variável fluxo de gás é a única variável controlável. As variáveis temperatura da água fria e caudal da água são variáveis não controláveis que dependem de factores externos.

A saída do sistema é a temperatura da água que sai do esquentador - temperatura da água quente  $t_{aq}(t)$ . Esta temperatura é função de todas as variáveis de entrada conhecidas  $t_{af}(t)$ ,  $c_a(t)$  e  $f_g(t)$  e de possíveis perturbações “desconhecidas”  $\xi(t)$ , como definido na equação 7.1:

$$t_{aq}(t) = f(t_{af}(t), c_a(t), f_g(t), \xi(t)) \quad (7.1)$$

As principais perturbações desconhecidas são o ruído de medida e as perdas de energia do sistema.

## 7.4 Gamas de Funcionamento

As gamas de funcionamento são as zonas de variação possíveis para as variáveis do sistema. Os limites inferiores e superiores são impostos por restrições físicas do equipamento ou pelas características intrínsecas das variáveis.

A tabela 7.1 mostra as gamas típicas de funcionamento das variáveis: temperatura da água quente, temperatura da água fria, fluxo de gás e caudal da água.

Por questões de segurança do utilizador, a temperatura da água quente é limitada superiormente a 65°C.

A gama da temperatura da água fria é dependente das condições climáticas da região do planeta e da estação do ano.

Variável	Acrón.	Gamas	Unid.
Temperatura da Água Quente	$t_{aq}$	[25; 65]	[°C]
Temperatura da Água Fria	$t_{af}$	[ 5; 30]	[°C]
Fluxo de Gás	$f_g$	[4; 18]	[l/min]
Caudal da Água	$c_a$	[3,5; 14]	[l/min]

Tabela 7.1: Gamas de funcionamento das principais variáveis.

O fluxo de gás depende da abertura da válvula proporcional, do tipo de gás utilizado e da respectiva pressão. A unidade da variável fluxo de gás, apresentada na tabela 7.1, é litros por minuto. Como a actuação no fluxo de gás é efectuada indirectamente através de um sinal do tipo PWM que define a corrente na válvula e a sua abertura, tornou-se mais prático utilizar a percentagem do factor de ciclo do sinal PWM em percentagem como unidade do fluxo de gás.

O esquema completo do sistema com a actuação na válvula de gás está ilustrado no diagrama de blocos da figura 7.6.

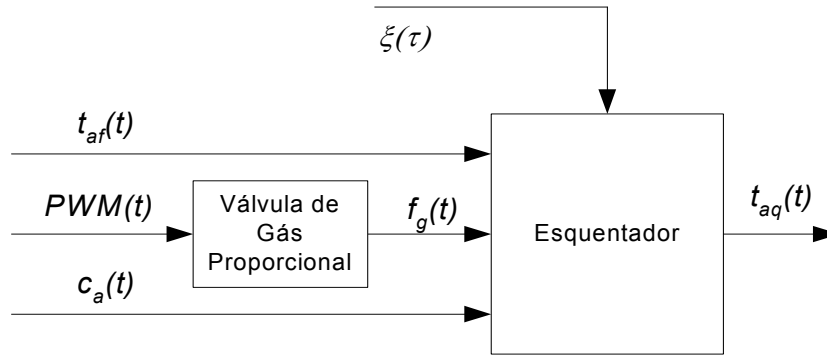


Figura 7.6: Modelo do esquentador com actuador incluído.

Os valores de fluxo de gás propano utilizado variam de 4 a 18 l/min<sup>1</sup>, correspondendo a valores de PWM variáveis entre 36 e 100%. Daqui em diante, a unidade utilizada para a variável fluxo de gás será a percentagem (%) que corresponde à percentagem de sinal PWM.

O fluxo de gás mínimo (36%) é imposto pelo sistema de modo a garantir uma combustão mínima, abaixo deste valor seria impossível manter a chama acesa no queimador. O fluxo de gás máximo (100%) é definido pela abertura máxima da válvula de gás e pela pressão máxima do gás fornecido.

O caudal da água máximo de funcionamento é limitado pelo sistema e controlado pelo utilizador e o caudal mínimo de funcionamento é imposto pelo sistema. Abaixo deste valor o esquentador não aquece a água.

<sup>1</sup>Valores medidos manualmente através de um sensor de pressão em forma de tubo graduado.

## 7.5 Estados de Funcionamento

O esquentador pode estar num dos cinco estados de funcionamento: desligado, arranque, operação, fecho e segurança. Estes estados são verificados e/ou impostos por um microcontrolador existente numa placa de controlo a apresentar com detalhe na secção 7.9.

No estado desligado não há circulação de água e os actuadores estão todos desligados. Neste estado, ao iniciar-se a circulação de água o esquentador passa para o estado de arranque.

O estado de arranque (ignição) ocorre sempre que existe água em circulação e não existe combustão de gás no queimador. Neste estado, o ventilador é ligado para fazer a exaustão dos gases de combustão, as válvulas de gás são abertas fornecendo gás ao queimador, o faiscador gera uma faísca junto do queimador e, finalmente, dá-se o início da combustão. Se alguns destes actuadores falharem ou se a combustão não se verificar, o sistema efectua uma segunda tentativa de arranque, efectuando novamente os passos descritos para o estado de arranque. Se após a segunda tentativa, não se verificar combustão no esquentador ou um dos actuadores não tiver ligado na devida altura, este passa para o estado de segurança.

No estado de segurança, todos os actuadores são desligados e o sistema é bloqueado, sendo impedido de arrancar novamente sem que antes um técnico verifique e corrija a causa desse bloqueio. Se o estado de arranque terminar com o início da combustão, então o esquentador passa para o estado de operação.

No estado de operação, é avaliado o desempenho de todos os actuadores. Se algum não estiver a funcionar correctamente, o esquentador passa para o estado de segurança. Se tudo estiver a funcionar correctamente, a água começa a ser aquecida devido à transferência de calor do ar aquecido pela combustão do gás para a água que circula dentro do permutador.

Logo que deixe de existir circulação de água no permutador, o esquentador passa ao estado de fecho. Neste, os actuadores são todos desligados com a excepção do ventilador, que ficará ainda a funcionar durante um período de tempo pré-definido (período de exaustão). Terminada a exaustão dos gases de combustão, o esquentador passa ao estado desligado.

O trabalho prático desenvolvido neste trabalho está centrado, essencialmente, no estado de operação, uma vez que é neste estado que o algoritmo de controlo da temperatura da água é executado.

## 7.6 Caracterização Estática e Dinâmica

A temperatura da água quente que sai do esquentador vai depender, essencialmente das três variáveis mensuráveis de entrada do sistema: temperatura da água fria, caudal de água, fluxo de gás e ainda, da dinâmica imposta pelas transferências de energias ocorridas no esquentador.

No estado de operação ocorrem três transferências de energia. Em primeiro lugar, ocorre a transferência de energia da combustão do gás para o ar contido dentro da

câmara de combustão. Em segundo, ocorre a transferência da energia do ar para o cobre do permutador. Por último, ocorre a transferência do cobre do permutador para a água que se encontra dentro do permutador. As constantes de tempo e os ganhos envolvidos nestas transferências de energia determinam o comportamento estático e dinâmico do esquentador.

Em estado estacionário, o esquentador é caracterizado por uma Potência Estática  $PE$  de funcionamento. Esta potência é a medida da energia fornecida à água por unidade de tempo para um determinado fluxo de gás constante. No caso presente é definida pela equação 7.2:

$$PE = (t_{aq} - t_{af}) c_a |_{f_g \in [36;100] \%} \quad [\text{Kcal/min}] \quad (7.2)$$

O esquentador é ainda caracterizado por uma potência estática máxima e mínima de funcionamento. Estas potências dependem, essencialmente, do tamanho físico do queimador, do número de injectores de gás, do tamanho da câmara de combustão e do permutador. A Potência Estática Máxima  $PE_{MAX}$  é a energia máxima por unidade de tempo que o esquentador consegue permutar em condições de segurança. Existem esquentadores de diferentes potências estáticas máximas, como por exemplo: 125, 275, 325 e 400 Kcal/min. Estes são valores de potência máxima típicos actualmente utilizados pelos vários fabricantes mundiais de esquentadores, como por exemplo a Vulcano Termodomésticos.

Quando é aplicado o fluxo de gás máximo na válvula de gás do esquentador, a potência estática máxima é definida pela equação 7.3:

$$PE_{MAX} = (t_{aq} - t_{af}) c_a |_{f_g=100 \%} \quad [\text{Kcal/min}] \quad (7.3)$$

Definindo a variável aumento de temperatura à saída do esquentador  $\Delta t$ , como sendo a subtracção entre as temperaturas da água quente e fria  $t_{aq} - t_{af}$ , a potência estática máxima pode ser redefinida pela equação 7.4:

$$PE_{MAX} = \Delta t c_a |_{f_g=100 \%} \quad [\text{Kcal/min}] \quad (7.4)$$

No caso presente, a potência estática máxima do esquentador é de 325 Kcal/min.

A potência estática mínima  $PE_{MIN}$  é a energia mínima por unidade de tempo que o esquentador permite permutar. Quando é aplicado o fluxo de gás mínimo constante na válvula de gás, a potência estática mínima é, no caso presente, definida pela equação 7.5:

$$PE_{MIN} = \Delta t c_a |_{f_g=36 \%} \quad [\text{Kcal/min}] \quad (7.5)$$

Este fluxo de gás garante uma combustão mínima assegurando que o esquentador se mantém no estado de operação (abaixo deste valor o esquentador a combustão não ocorre passando para o estado de desligado). A potência estática mínima de um esquentador pode variar de 70 a 100 Kcal/min, sendo de 80 Kcal/min no esquentador utilizado.

Desenvolvendo a equação 7.2 em ordem à temperatura da água quente, verifica-se que esta é igual ao quociente entre a potência estática e o caudal da água mais a temperatura da água fria, como definido na equação 7.6:



$$t_{aq} = \frac{PE}{c_a} + t_{af} \quad [^{\circ}\text{C}] \quad (7.6)$$

Analisando a influência da temperatura da água fria verifica-se que as variações desta são muito lentas, uma vez que dependem, fundamentalmente, das condições climáticas, da estação do ano e da região do planeta. Em cada período de operação do esquentador as variações da temperatura da água fria são, praticamente nulas permanecendo praticamente constante em cada período de operação.

Ao nível estático, a temperatura da água fria provoca, simplesmente, como definido na equação 7.6, um desvio na temperatura da água quente. Ao nível dinâmico, a variação da temperatura da água quente (variação de  $t_{aq}$ ) é equivalente à variação do aumento da temperatura da água (variação de  $\Delta t$ ), uma vez que, se parte do princípio de que a temperatura da água fria permanece praticamente constante em cada período de operação.

Daqui em diante, em vez de utilizar como variável de saída do esquentador a temperatura da água quente, utiliza-se o aumento da temperatura da água, simplificando a representação do sistema e “eliminando” uma das variáveis de entrada.

O diagrama de blocos do sistema simplificado, apresentado na figura 7.7, passa a ter duas variáveis de entrada : caudal da água  $c_a$  e fluxo de gás  $f_g$  e uma variável de saída: aumento de temperatura da água quente  $\Delta t$ . A variável  $\xi$  é uma variável não mensurável que representa as perturbações.

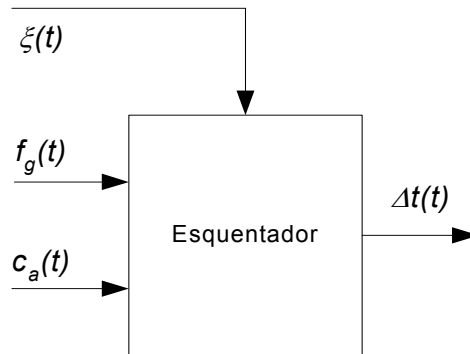


Figura 7.7: Variáveis de entrada e de saída do esquentador (modelo simplificado).

O estudo do comportamento estático e dinâmico do esquentador é apresentado de seguida sendo baseado na observação experimental e na análise das relações existentes entre as variáveis de entrada e saída do sistema.

### 7.6.1 Funcionamento Estático

#### Relação entre Fluxo de Gás, Caudal da Água e Aumento da Temperatura

Todos os pontos de funcionamento do esquentador são caracterizados por uma potência estática de funcionamento, como definido na equação 7.2.

Com o objectivo de conhecer a superfície dos pontos de funcionamento estático, aplicaram-se fluxos de gás desde o valor mínimo até ao valor máximo para caudais de água desde o valor mínimo até ao valor máximo e mediram-se os respectivos aumentos, em estado estacionário, da temperatura à saída do sistema. Desta sequência de testes efectuada resultou a superfície de pontos de funcionamento estático ilustrada na figura 7.8.

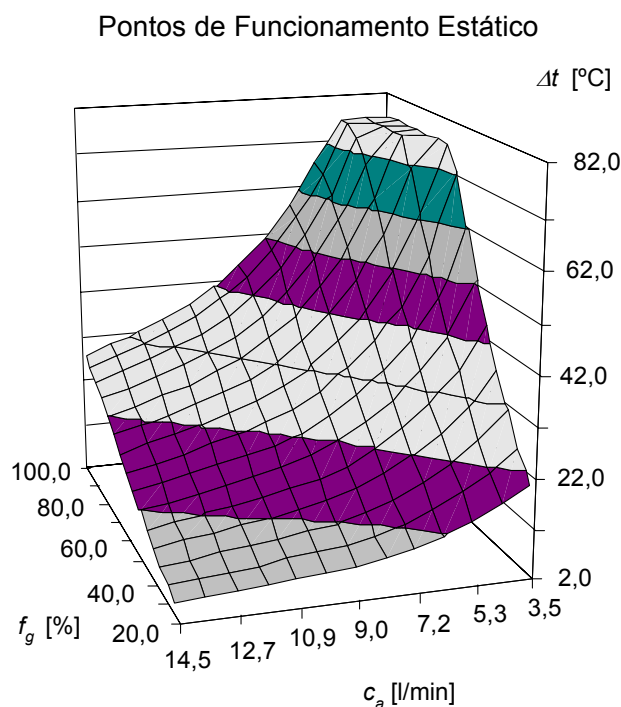


Figura 7.8: Superfície com os pontos de funcionamento estático.

Da análise da figura 7.8, verifica-se que as duas variáveis de entrada (fluxo de gás e caudal da água) alteram fortemente o ponto de funcionamento do sistema.

A relação entre o aumento da temperatura e o caudal de água, como era previsível através da equação 7.2, é “fortemente” não linear uma vez que, ao nível estático, o aumento da temperatura é inversamente proporcional ao caudal da água.

Através da figura 7.8, observa-se ainda que a relação entre o fluxo de gás e o aumento de temperatura é não linear. Contudo, esta não linearidade é menos acentuada do que a verificada entre as variáveis caudal de água e aumento de temperatura. Para verificar esta não linearidade mais detalhadamente, aplicaram-se para um caudal de água fixo incrementos lineares de fluxo de gás e registaram-se os aumentos de temperatura em estado estacionário. O resultado destes testes pode ser observado na figura 7.9.

A figura mostra a existência de uma relação “ligeiramente” não linear entre o aumento de temperatura e o fluxo de gás. Pode observar-se ainda que não existe histerese entre o aumento da temperatura e o fluxo de gás, uma vez que, a diferença entre a curva ascendente e a curva descendente do aumento da temperatura é praticamente nula.

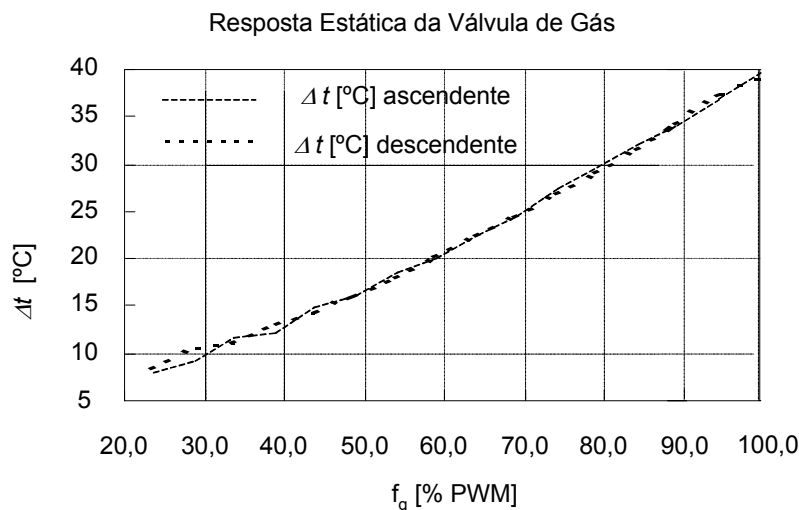


Figura 7.9: Aumentos da temperatura em função do fluxo de gás, para caudal fixo de 8,5 l/min. (respostas em estado estacionário).

### 7.6.2 Funcionamento Dinâmico

O estudo e análise do comportamento dinâmico do sistema baseou-se na observação da dinâmica das relações entre as variáveis de entrada fluxo de gás e caudal da água, e a variável de saída aumento da temperatura.

Neste sentido, foi efectuada a aquisição de dados de entrada e saída do sistema com um período de 250 ms. Este período de amostragem permitiu uma observação detalhada da evolução dinâmica das várias variáveis ao longo do tempo.

#### Relação entre Fluxo de Gás e o Aumento da Temperatura

Para o estudo da relação dinâmica entre fluxo de gás e aumento da temperatura, aplicaram-se vários sinais do tipo degrau ao fluxo de gás e mediram-se os tempos de subida e descida do aumento da temperatura<sup>2</sup>. Estes testes foram efectuados para vários caudais.

Na figura 7.10 estão ilustrados os testes efectuados: com caudal reduzido (5,5 l/min) e com caudal elevado (10,8 l/min). Estes testes mostram que para um determinado caudal fixo, os tempos de subida e descida ( $\tau_s$  e  $\tau_d$ ) são iguais e ambos diminuem com o aumento do caudal.

Observando a figura 7.10 verifica-se que, para um determinado caudal fixo, o aumento da temperatura tem uma evolução do tipo exponencial mostrando assim que o esquentador apresenta uma dinâmica semelhante à dinâmica de um sistema de primeira ordem.

<sup>2</sup> Apesar dos tempos de subida e descida serem grandezas utilizadas na caracterização de sistemas lineares, quando usados em sistemas não lineares, estes fornecem uma “boa indicação” sobre a dinâmica do sistema.

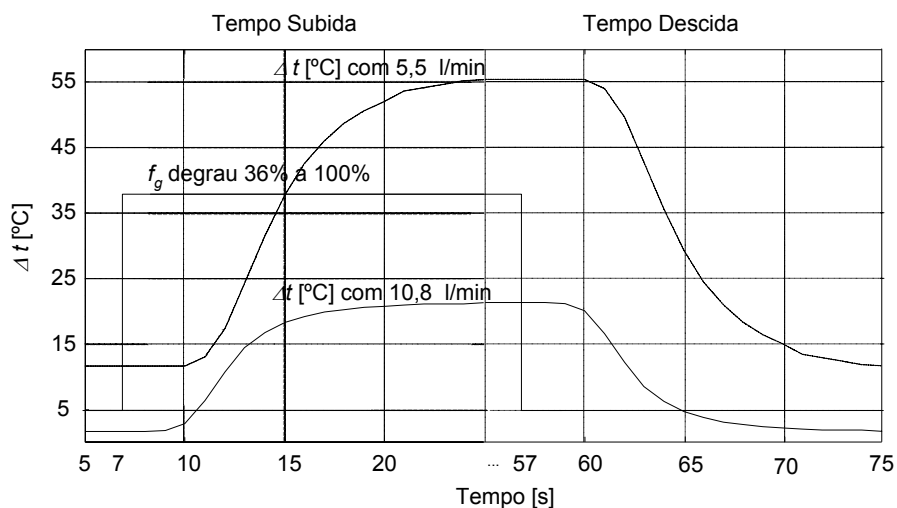


Figura 7.10: Aumentos da temperatura em função ao degrau ascendente e descendente de fluxo de gás para dois caudais diferentes.

Para cinco caudais diferentes e aplicando sinais do tipo degrau ao fluxo de gás obtiveram-se os resultados ilustrados na figura 7.11. Desta figura verifica-se que existe um tempo morto  $\tau_{m1}$  variável entre o fluxo de gás e aumento da temperatura que é dependente do caudal da água.

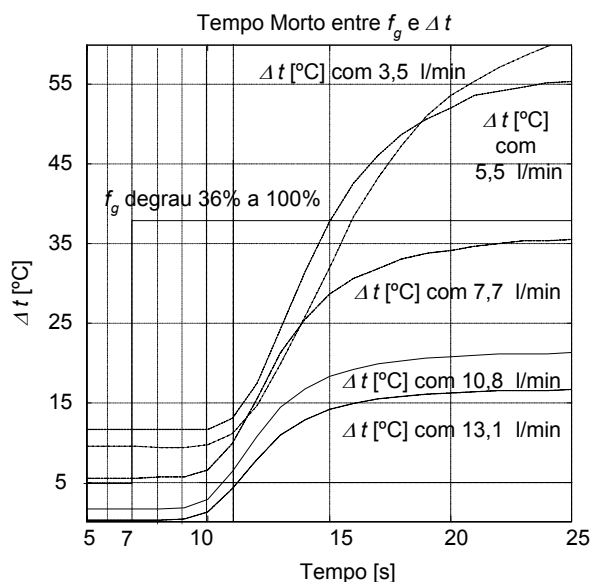


Figura 7.11: Aumentos da temperatura da água em função ao degrau ascendente de fluxo de gás para cinco caudais fixos diferentes (pormenor do tempo morto).

Ainda com base na figura 7.11 construiu-se a tabela 7.2 que apresenta um sumário do tempo morto, dos tempos de subida e descida e da variação do aumento da temperatura

para os cinco caudais.

<b>Caudal (fixo)</b>	<b>3,5</b>	<b>5,5</b>	<b>7,7</b>	<b>10,8</b>	<b>13,1</b>	<b>[l/min]</b>
Tempo Morto	3,75	3,50	3,25	3,00	2,75	[s]
Tempo Subida	19,25	11,00	6,75	5,25	5,00	[s]
Tempo Descida	19,25	11,00	6,75	5,25	5,00	[s]
$\partial(\Delta t) = \Delta t_{final} - \Delta t_{inicial}$	55,3	46,5	32,2	19,1	16,0	[°C]

Tabela 7.2: Resumo das respostas do aumento da temperatura a degraus de fluxo de gás para cinco caudais diferentes.

Na tabela 7.2 observa-se que os tempos de subida e descida são iguais para um determinado caudal fixo. Observa-se ainda que a variação de aumento de temperatura  $\partial(\Delta t)$  cresce com a diminuição do caudal. Isto permite afirmar que o esquentador possui uma “constante de tempo” e um “ganho” que variam com o caudal da água.

O esquentador apresenta um tempo morto variável entre o fluxo de gás e o aumento da temperatura que depende essencialmente do caudal de água. Analisando mais detalhadamente a variação do tempo morto, verifica-se que este é composto por uma componente fixa e outra variável. A componente fixa surge do tempo dispendido desde da abertura da válvula de gás até ao início do aquecimento do permutador. A componente variável surge do tempo dispendido pela água em percorrer o tubo até atingir o sensor de temperatura da água quente. Durante este percurso não existe transferência de energia para a água.

O tempo morto total  $\tau_{m1}(t)$  está definido na equação 7.7:

$$\tau_{m1}(t) = \tau_{m0}(t) + \tau_{m2}(t) = C_1 + \pi r^2 l \frac{C_2}{c_a(t)} \quad [\text{s}] \quad (7.7)$$

em que  $C_1$  é a componente fixa do tempo morto,  $r$  é o raio do tubo,  $l$  é o comprimento do tubo e  $C_2$  é o factor de correcção de escala. A componente fixa do tempo morto  $\tau_{m0}$  é aproximadamente igual 2,50 s sendo este o tempo necessário para que o gás chegue ao queimador, seja queimado e comece a aquecer o permutador. Este valor foi medido através de testes experimentais efectuados no esquentador registando a temperatura do permutador.

A segunda componente do tempo morto  $\tau_{m2}$  é calculada através do caudal da água. O número de litros por unidade de tempo é equivalente a um volume ( $V$ ) de água por unidade de tempo. Através do volume da água a transportar é possível saber quanto tempo demora esse volume de água a ser totalmente transportado.

A expressão do volume é dada pela equação 7.8:

$$V = \pi r^2 l \quad [\text{m}^3] \Leftrightarrow V * 1000 \quad [\text{litros}] \quad (7.8)$$

e sendo o respectivo caudal de água dado pela equação 7.9:

$$c_a \quad [\text{litros/segundo}] \Leftrightarrow \frac{c_a}{60} \quad [\text{litros/s}] \quad (7.9)$$

Conclui-se das equações anteriores que o tempo  $\tau_{m2}$  necessário para transportar o volume  $V$  de litros até à saída, é dado pela equação 7.10:

$$\tau_{m2} = \frac{1000V}{\frac{ca}{60}} = \pi r^2 l \frac{60000}{C_a} \text{ [s]} \quad (7.10)$$

Na figura 7.12, está representado o tempo morto “teórico” (segundo a equação 7.7) e o tempo morto prático experimental (segundo a tabela 7.2) em função do caudal.

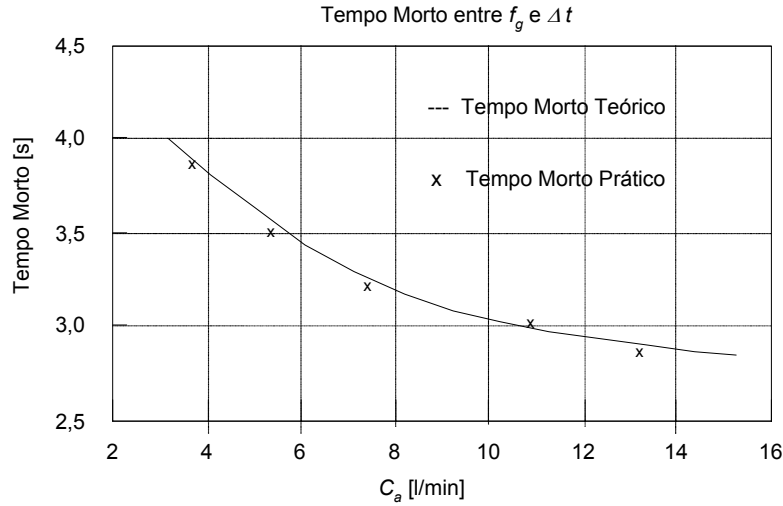


Figura 7.12: Tempo morto teórico e tempo morto prático em função do caudal.

Da análise da figura 7.12 pode-se afirmar que a equação 7.7 fornece uma boa aproximação do tempo morto entre o fluxo de gás e o aumento da temperatura.

### Relação entre Caudal da Água e o Aumento da Temperatura

Para o estudo da relação dinâmica entre as variáveis caudal da água e aumento da temperatura, aplicaram-se vários sinais do tipo degrau no caudal da água e mediram-se os tempos de subida e descida da variável aumento de temperatura  $\Delta t$ , utilizando um fluxo de gás constante próximo de 65% (valor intermédio da gama de funcionamento do fluxo de gás).

Dos testes efectuados um deles está ilustrado na figura 7.13 no qual se pode observar que os tempos de subida e descida são muito diferentes para degraus simétricos efectuados no caudal.

Na figura 7.14 estão ilustrados mais dois dos testes efectuados onde se observa a existência de um tempo morto variável entre o caudal da água e o aumento da temperatura, como apresentado na equação 7.7 (componente variável). Este tempo morto corresponde ao tempo do transporte da água até atingir o sensor de água quente.

Ainda através da figura 7.14 pode observar-se que, para uma variação em degrau do caudal da água de 12,5 l/min para 8,5 l/min o tempo morto é cerca de 0,5 s e para uma variação em degrau do caudal da água de 12,5 l/min para 5,5 l/min o tempo morto é cerca de 1 s.

Para variações do caudal de água do mínimo (3,5 l/min) para o máximo (14,0 l/min) e do máximo para o mínimo, observam-se, respectivamente, tempos mortos de 0,25 s e

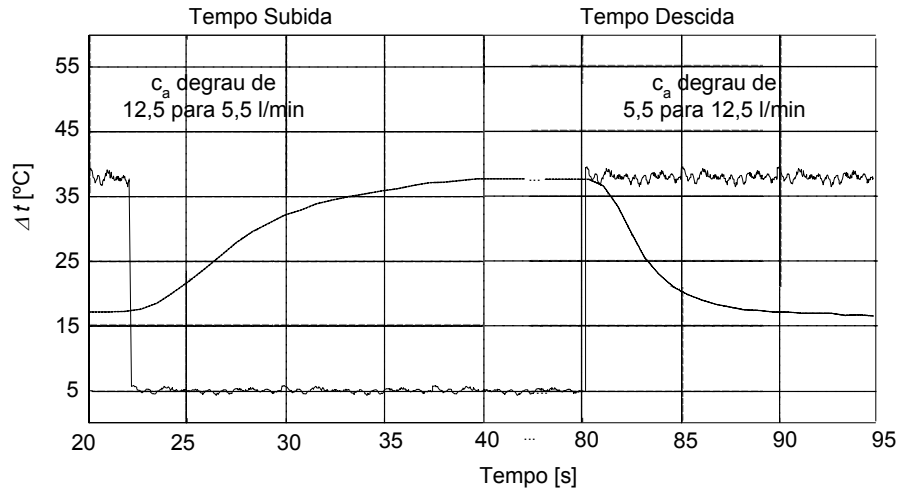


Figura 7.13: Aumento da temperatura em função de degraus ascendente e descendente de caudal da água, para um fluxo de gás fixo de 65%.

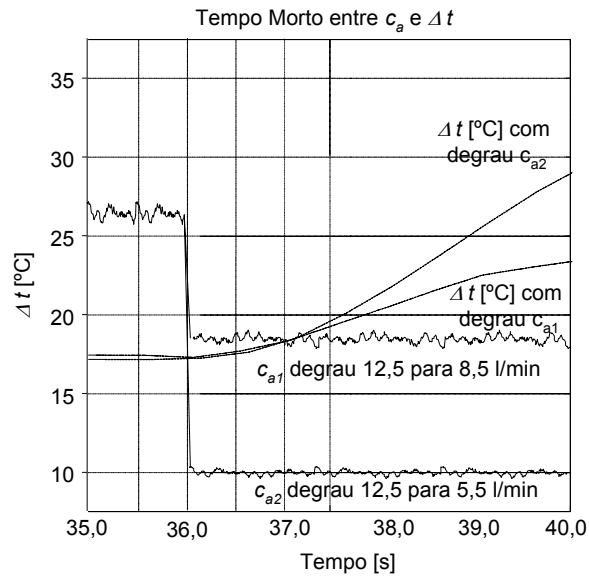


Figura 7.14: Aumento da temperatura em função de degraus descendentes de dois caudais da água, para um fluxo de gás fixo de 65% (pormenor do tempo morto).

1,25 s. Estes valores estão muito próximos dos valores calculados teoricamente através da equação 7.10 e correspondem ao tempo morto variável do sistema  $\tau_{m2}(t)$ , entre o caudal de água e o aumento da temperatura.

### 7.6.3 Comentários

Com base na caracterização estática e dinâmica efectuada ao esquentador, pode concluir-se que este possui uma “constante de tempo” dominante e um “ganho” que variam com o caudal da água. A constante de tempo surge da transferência de energia do cobre do permutador para a água que se encontra no seu interior. As restantes transferências de energia ocorrem muito mais rapidamente do que a anterior, pelo que se desprezaram as respectivas constantes de tempo.

Verifica-se ainda que o esquentador possui dois tempos mortos. O tempo morto existente entre o fluxo de gás e o aumento da temperatura é composto por uma componente fixa e uma componente variável. A sua gama variação está contida no intervalo  $[2,75; 3,75]$  [s]. O tempo morto existente entre o caudal da água e o aumento da temperatura é constituído só pela componente variável, estando a sua gama de variação contida no intervalo  $[0,25; 1,25]$  [s].

Em forma de resumo a figura 7.15 apresenta o diagrama de blocos do esquentador e as suas características: os componentes físicos, as variáveis de entrada e saída, as não linearidades e os respectivos tempos mortos.

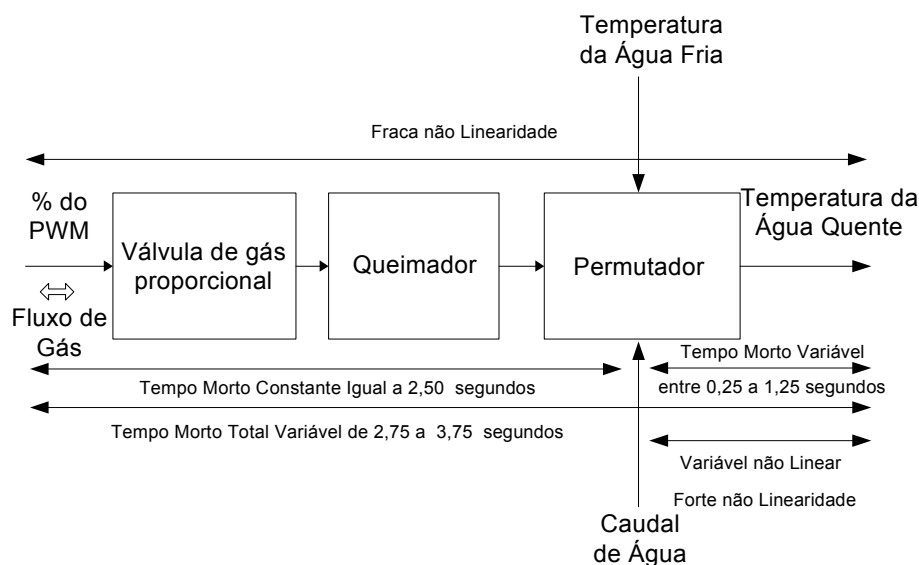


Figura 7.15: Diagrama de blocos do esquentador e suas características funcionais.

## 7.7 Especificações de Desempenho

As especificações de desempenho para o controlo automático do esquentador utilizadas neste trabalho, foram baseadas nas especificações definidas por um dos grandes fabricantes de esquentadores português e europeu: a Vulcano Termodomésticos.

As especificações exigidas ao controlo do esquentador doméstico são:



- A diferença entre a temperatura da água de saída e a temperatura da água desejada (erro em estado estacionário) deve estar contida no intervalo  $[-1,0; +1,0]$   $[^{\circ}\text{C}]^3$ ;
- O tempo de estabelecimento da resposta da temperatura da água quente para variações da temperatura desejada e/ou do caudal deve estar contido no intervalo  $[0; 15]$  [s].
- A temperatura da água desejada é limitada pelo sistema ao intervalo  $[30,0; 60,0]$   $[^{\circ}\text{C}]$ .

O sistema de controlo desenvolvido tem de respeitar estas especificações em qualquer situação, em estado transitório e/ou estado estacionário.

No caso em que o utilizador peça uma potência estática maior do que a potência estática máxima do esquentador, este deve fornecer a sua potência máxima. No caso em que o utilizador peça uma potência estática menor do que a potência estática mínima do esquentador, este deve fornecer a sua potência mínima. Excepcionalmente, para estas duas situações as especificações da malha de controlo deixam de ser exigidas.

## 7.8 Plataforma de Testes

Uma vez que este trabalho exigiu o teste de diversos algoritmos de controlo baseados em modelos, tornou-se necessária a implementação de uma plataforma de testes que permitisse, de uma forma simples e expedita, a realização dos referidos testes. Esta plataforma permitiu ainda a aquisição dos dados de entrada e saída do sistema para o treino e teste dos vários modelos propostos e a monitorização do desempenho dos vários algoritmos de controlo propostos.

A plataforma de testes implementada é constituída, como ilustrado na figura 7.16, por um esquentador, uma placa de controlo (baseada num microcontrolador) e um computador pessoal.

Na fase de desenvolvimento os vários algoritmos de controlo não foram executados pelo microcontrolador da placa de controlo, mas sim pelo computador pessoal. Este procedimento simplificou, em grande medida, a avaliação e o teste dos vários algoritmos de controlo, não necessitando de reprogramar o microcontrolador a cada novo teste. Só na fase final é que os algoritmos de controlo foram implementados e executados no microcontrolador 89C51RD.

Os modelos e os algoritmos de controlo desenvolvidos neste trabalho foram executados no computador pessoal no ambiente MATLAB [Mat99]. Antes de testar os algoritmos de controlo no sistema real, estes foram testados previamente em simulação utilizando o modelo do esquentador como substituto do esquentador real.

O ambiente MATLAB foi escolhido para a definição e identificação dos modelos e simulação dos algoritmos de controlo desenvolvidos, uma vez que é um ambiente de

---

<sup>3</sup>Esta especificação refere-se ao erro em estado estacionário que mesmo na presença de ruído nas medidas deve estar contido no referido intervalo.

Associado a esta especificação é exigido ainda ao controlador que as sobrelevações/subelevações relativas às variações do sinal de referência sejam inferiores a  $1^{\circ}\text{C}$  em modulo.

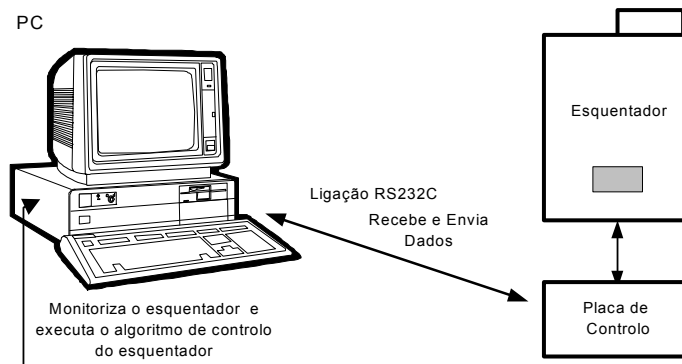


Figura 7.16: Esquema global da plataforma de testes.

trabalho que possui uma grande quantidade de ferramentas de *software* já desenvolvidas, como as ferramentas de modelização, identificação, controlo e muitas outras, tornando o desenvolvimento de novas ferramentas muito mais rápido e simples. Além das vantagens referidas, este ambiente facilita o desenvolvimento de novas ferramentas, uma vez que possui uma linguagem de programação simples e orientada para o cálculo numérico, especificamente, vocacionado para a manipulação de matrizes.

A ligação deste ambiente de trabalho à placa de controlo acabou por se mostrar de grande importância, uma vez que, não só permite a aquisição de dados de entrada e saída do sistema, para o treino e validação dos modelos, como também o teste dos algoritmos de controlo desenvolvidos de uma forma simples e expedita. O diagrama de blocos da plataforma de testes desenvolvida está ilustrado da figura 7.17.

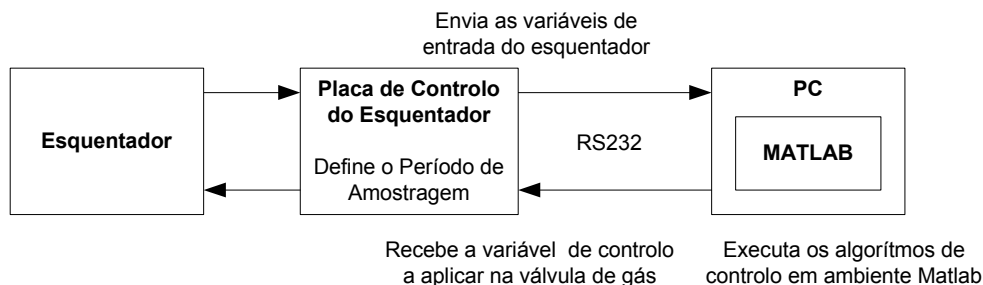


Figura 7.17: Diagrama de blocos das comunicações entre a placa de controlo e o computador pessoal.

Para estabelecer as comunicações entre o ambiente MATLAB e o microcontrolador da placa de controlo, utilizou-se a *Real Time Toolbox 3.0* para MATLAB [Hum99] que permite a comunicação entre o MATLAB e a porta série do computador.

O microcontrolador da placa de controlo é responsável pelo envio das variáveis de entrada do esquentador para o computador pessoal através de uma linha série RS232. No computador pessoal, mais concretamente no ambiente MATLAB, as variáveis de entrada são recebidas, o algoritmo de controlo é executado e a nova variável de controlo

é enviada para a respectiva placa. O microcontrolador desta placa perante a nova variável de controlo actualiza os actuadores do esquentador ficando a aguardar até à próxima amostragem.

Durante o teste e desenvolvimento dos algoritmos de controlo para o esquentador, apresentados neste trabalho, a placa de controlo funcionou, fundamentalmente, como uma placa de interface entre o esquentador e o computador pessoal. Para a construção do protótipo final de um esquentador controlado automaticamente, esta placa, além das funções já referidas anteriormente, passa a executar o algoritmo de controlo, dispensando a presença do computador pessoal.

A figura 7.18 mostra uma fotografia da bancada de testes utilizada durante o decorrer deste trabalho, constituída pelo esquentador, pela placa de controlo e pelo computador pessoal.

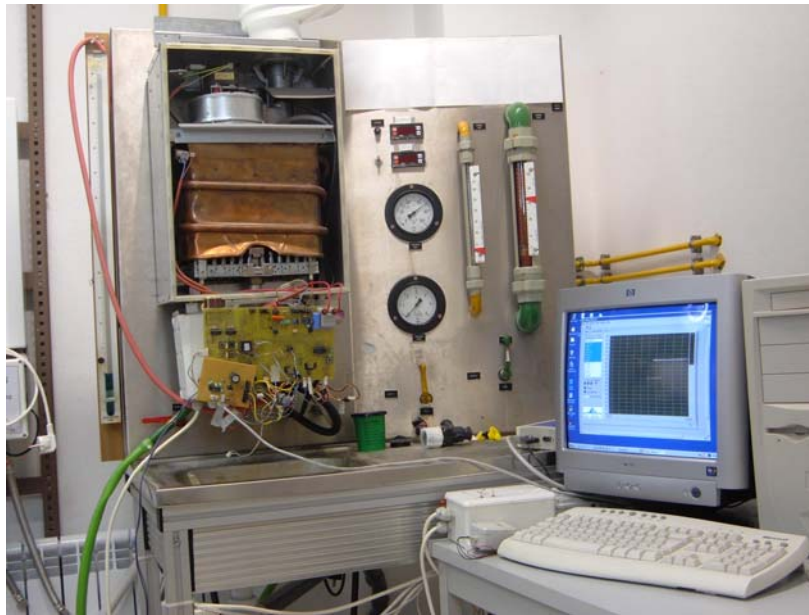


Figura 7.18: Fotografia da plataforma de testes.

## 7.9 Placa de Controlo

A placa de controlo é constituída por um microcontrolador principal 89C51RD da PHILIPS, por circuitos electrónicos de condicionamento de sinais dos sensores e dos actuadores, por circuitos de interface com o utilizador, por um segundo microcontrolador responsável pela segurança do sistema e, finalmente, por circuitos necessários para o estabelecimento de comunicações bilaterais entre a placa e o computador pessoal. A figura 7.19, mostra uma fotografia de toda a electrónica integrada numa placa de circuito impresso. Esta placa é colocada na parte inferior do esquentador tornando-se parte integrante deste, como ilustrado na figura 7.18.

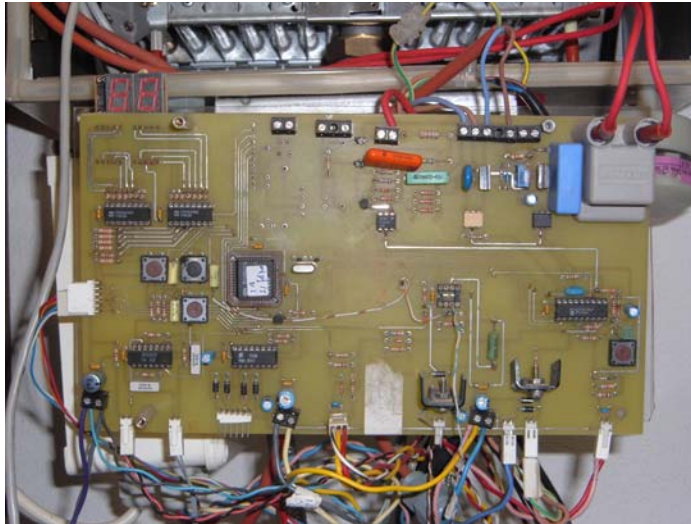


Figura 7.19: Fotografia da placa de controlo.

O projecto e a escolha dos sensores, actuadores, circuitos electrónicos de condicionamento de sinal e dos circuitos de interface com os actuadores, foram definidos e impostos no início do projecto pela empresa TetraCis - Electrónica e Telecomunicações, Lda, assim como, o posicionamento dos vários sensores no sistema. Deste modo, não foi possível efectuar qualquer alteração no *hardware* que pudessem melhorar e ou simplificar a aquisição e o controlo das variáveis do sistema. O circuito eléctrico completo da placa de controlo pode ser observado no Apêndice C - Sistema de Controlo do Esquentador - *Hardware*.

Os cinco módulos fundamentais da placa de controlo são:

- Microcontrolador
- Interface com sensores e actuadores
- Interface com o utilizador
- Comunicações
- Segurança

O diagrama de blocos da figura 7.20 ilustra os referidos blocos e sua interligação com o microcontrolador.

O microcontrolador é responsável pela definição do período de amostragem  $h$  e pela interligação/controlo de todos os módulos existentes na referida placa. O microcontrolador principal só inicia o arranque do sistema quando o microcontrolador secundário existente no módulo de segurança, lhe der a respectiva autorização. Quando o esquentador se encontra no estado de operação, o microcontrolador principal lê os sinais referentes às variáveis de entrada, estabelece as comunicações com o computador pessoal, recebe o sinal de actuação, controla os vários actuadores, e por último, mostra num *display* de dois dígitos a temperatura da água à saída do esquentador.

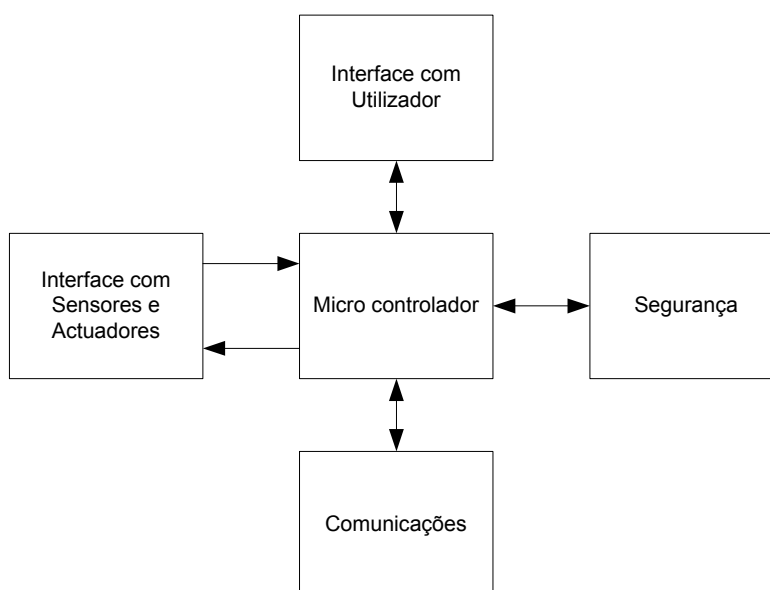


Figura 7.20: Diagrama de blocos da placa de controlo do esquentador.

O módulo de interface com os sensores e actuadores é constituído pelos circuitos de condicionamento de sinal, permitindo a leitura das variáveis de temperatura de água quente e fria, do caudal de água, de ionização, de excesso de temperatura e de ventilação, e pelos circuitos de interface com os actuadores, permitindo a actuação e controlo do faiscador, ventilador e válvulas de gás proporcional e *on/off*.

Os sensores que lêem as variáveis de entrada e saída do sistema são: duas resistências de coeficiente negativo (NTC) e um sensor de turbina, permitindo medir, respectivamente, as temperaturas da água fria, quente e o caudal da água. As resistências de coeficiente negativo são ligadas a circuitos osciladores, gerando sinais do tipo onda quadrada cujo período é inversamente proporcional ao valor das respectivas resistências. O microcontrolador, através de um contador de 16 *bits*, mede os períodos, permitindo obter as duas temperaturas. Devido à não linearidade das NTCs a variação destes períodos é não linear com as temperaturas sendo a linearização destas realizada por *software*. As características das NTCs e os métodos de conversão utilizados permitem a medição de temperaturas com uma resolução de décimas de grau centígrado.

O sensor de turbina contém um dispositivo óptico linear que gera um sinal eléctrico do tipo onda quadrada cuja frequência é proporcional ao caudal. O microcontrolador, através de um contador de 8 *bits*, mede o respectivo período que é inversamente proporcional ao caudal de água. As características do sensor e o método de conversão utilizados permitem a medição de caudais com uma resolução de décimas de litro por minuto.

O actuador principal do sistema é uma válvula de gás proporcional controlada por corrente contínua. Esta corrente é função do sinal modulado por largura de impulso (PWM) aplicado a um circuito *on/off* baseado num Mosfet (ver Anexo C), permitindo controlar a abertura da válvula. Este sinal é gerado pelo microcontrolador através de

um registo de 8 *bits*, permitindo uma actuação com uma resolução de 256 níveis.

Os restantes sensores e actuadores funcionam de modo binário com circuitos de interface muito simples e conhecidos (ver Anexo C).

O módulo de interface com o utilizador permite a visualização a temperatura da água quente, temperatura da água desejada e os códigos de erros provenientes do estado de segurança.

O sistema possui dois dígitos de sete segmentos e dois botões de pressão que proporcionam o interface de comunicação entre a placa de controlo e o utilizador.

Quando o esquentador se encontra em operação os dois dígitos mostram a temperatura da água quente. Quando o erro entre a temperatura desejada e da temperatura à saída estiver fora do intervalo  $[-3.0; +3.0]$  [°C], os dois dígitos mostram a temperatura de forma intermitente de maneira a alertar o utilizador da existência desse erro. No estado desligado, os dois dígitos mostram a última temperatura de água desejada definida pelo utilizador. No estado de segurança, os dois dígitos mostram um de vários símbolos, correspondendo a um de vários tipos de erros possíveis, de forma intermitente.

A placa de controlo possui ainda dois botões de pressão que permitem incrementar ou decrementar a temperatura da água desejada. Durante este processo é dada a conhecer a temperatura desejada definida pelo utilizador.

O módulo de comunicações é responsável pelo estabelecimento da comunicação bi-direccional entre o microcontrolador e o computador pessoal através de uma linha série RS232C. O computador pessoal recebe os valores numéricos das variáveis de entrada do esquentador, calcula o sinal de controlo e envia-o para o microcontrolador que, por sua vez, actua no sistema. A comunicação entre a placa de controlo e o computador pessoal foi descrita com mais detalhe na secção anterior.

O módulo de segurança é responsável pela segurança do esquentador de uma forma permanente, dando particular atenção aos estados de arranque, operação e fecho. Este módulo é executado pelo microcontrolador PIC 16F84 da MICROCHIP, que supervisiona os sensores de ionização, excesso de temperatura, ventilação e controla o faiscador, o ventilador e a válvula de gás *on/off*.

O módulo de segurança é responsável pela actualização do estado actual de funcionamento do esquentador. Este estado é definido com base na leitura dos vários sensores existentes no sistema. Se todas as variáveis estiverem dentro das suas gamas de funcionamento e dos padrões de segurança predefinidos, o esquentador funciona normalmente, caso contrário, este módulo bloqueia o esquentador impedindo-o de continuar a funcionar.

## 7.10 Conclusão

Este capítulo apresentou uma descrição pormenorizada das características físicas e funcionais de um esquentador doméstico, os requisitos de controlo exigidos para o sistema, a plataforma de testes implementada para a validação dos vários algoritmos

de controlo desenvolvidos e, finalmente, descreveu a placa de controlo utilizada no controlo de todo o sistema.

Na secção 7.2, efectuou-se uma descrição dos vários elementos físicos que constituem um esquentador. Na secção 7.3, enumeraram-se as variáveis de entrada e saída do esquentador. As zonas de funcionamento e os limites mínimos e máximos das variáveis do sistema foram definidos na secção 7.4. Na secção 7.5, foram descritos os vários estados de funcionamento do equipamento.

O estudo do comportamento estático e dinâmico do esquentador foi apresentado na secção 7.6. Este estudo foi baseado em testes experimentais permitindo a observação das relações estáticas e dinâmicas entre as variáveis de entrada e saída.

Na secção 7.7, definiram-se as especificações de desempenho exigidas para o controlo do esquentador. Na secção 7.8, foi apresentada a plataforma de testes sendo constituída por um computador pessoal, uma placa de controlo e um esquentador. Na secção 7.9, descreveu-se a placa de controlo desenhada para o controlo de todo o sistema.

A caracterização do esquentador doméstico permitiu concluir que este apresenta um comportamento aproximado ao de um sistema de primeira ordem de parâmetros variáveis com tempo morto relativamente grande e variável. Estes parâmetros variáveis dependem não linearmente do caudal de água e do fluxo de gás. Esta caracterização permitiu ainda a aquisição de conhecimentos experimentais sobre o sistema determinantes para na construção de modelos e controladores para o esquentador propostos nos capítulos seguintes.





## Capítulo 8

# Modelização e Identificação do Esquentador

### 8.1 Introdução

Este capítulo apresenta os resultados da aplicação dos vários métodos de modelização do tipo caixa cinzenta propostos nos capítulos anteriores na modelização de um esquentador doméstico. É apresentado o processo de aquisição de dados, é definido o modelo paramétrico do esquentador, as estruturas de treino para o modelo directo e inverso e, finalmente, são definidos, identificados e avaliados os modelos directos e inversos: linear, neuro-difuso, Hammerstein adaptativo polinomial/neuro-difuso e híbrido série polinomial/neuro-difuso. Todos os modelos apresentados são modelos sem tempo morto, sendo este último modelizado num bloco à parte. O modelo linear é apresentado apenas para comparação dos resultados obtidos com os restantes modelos não lineares.

Ao nível dos modelos neuro-difusos, foram testadas várias abordagens utilizando diferentes regressores [VM02], com diferentes ferramentas de construção e identificação [VM03b] e utilizando diferentes métodos de inicialização das respectivas estruturas [VM03a] permitindo apresentar nesta dissertação um modelo neuro-difuso com um número mínimo de regressores, moderada complexidade, baixo erro de modelização e de rápida velocidade de convergência.

Com base na modelização orientada por blocos, foram construídos e identificados dois modelos Hammerstein adaptativos, um que utiliza funções polinomiais e outro que utiliza sistemas neuro-difusos para a modelização da função não linear estática [VM04c] e [VM04a].

Ao nível da modelização híbrida foram propostos dois modelos híbridos série, um que usa funções polinomiais e outro que usa sistemas neuro-difusos para a modelização da componente desconhecida do modelo físico [VM04f] e [VM05].

Finalmente, são apresentadas as conclusões, comparando os erros de modelização, a interpretabilidade e a complexidade matemática dos vários modelos do tipo caixa cinzenta propostos.

Neste capítulo encontram-se algumas das contribuições apresentadas neste trabalho. A proposta de definição da equação de balanço de energia para um esquentador

doméstico genérico, a definição e identificação dos modelos Hammerstein adaptativos polinomiais e neuro-difusos e a definição e identificação dos modelos híbridos polinomiais e neuro-difusos.

## 8.2 Aquisição de Dados

Os dados de entrada e saída são umas das fontes de conhecimento mais utilizadas na modelização e identificação de sistemas. Mesmo quando se utilizam outros tipos de conhecimentos para a construção de modelos, existem sempre parâmetros dos modelos que são desconhecidos e que são identificados recorrendo aos dados de entrada e saída. Estes dados são ainda importantes para o teste e validação dos modelos, uma vez que permitem estudar a qualidade e capacidade de generalização dos mesmos.

Como referido no capítulo 2, a recolha de dados é uma das primeiras etapas a realizar no processo de identificação. Assim, esta etapa comporta a escolha do período de amostragem e ainda a aquisição dos dados para a identificação.

### 8.2.1 Escolha do Período de Amostragem

A escolha do período de amostragem é, como referido no capítulo 2, um passo importante na qualidade dos dados recolhidos.

Do estudo preliminar do esquentador efectuado no capítulo 7 conclui-se que, globalmente, o esquentador possui um ganho e uma constante de tempo dominante que variam fortemente com o caudal da água. Apesar da não linearidade verificada no sistema, a definição do período de amostragem foi baseada na relação definida na equação 8.1, e nos valores obtidos para os tempos de subida do esquentador apresentados no capítulo 7:

$$h = \frac{\tau_s}{N_r} \text{ em que } N_r \simeq 4 \text{ a } 10 \quad (8.1)$$

em que  $\tau_s$  é o tempo de subida,  $h$  é o período de amostragem e  $N_r$  é o número de amostras contidas no tempo de subida.

Do estudo preliminar apresentado no capítulo 7, verifica-se que os valores do tempo de subida variam de 19,25 s a 5,00 s para caudais de 3,5 a 13,1 l/min, respectivamente. Deste modo, o período de amostragem mínimo que respeita a regra definida na equação 8.1 é o período de um segundo. Com este período de amostragem é garantido um número mínimo de amostras recolhidas para o menor tempo de subida.

A definição do período de amostragem também é dependente do tempo morto do sistema. O período de amostragem deve ser, pelo menos, duas a três vezes menor do que o tempo morto do sistema [SEM89] e [DC03]. O período de amostragem escolhido para o esquentador, também respeita esta regra como se poderá ver mais adiante neste capítulo.

O conhecimento físico da variação do tempo de subida com o caudal sugere a utilização de um período de amostragem variável. Este foi um dos pontos que ficou em aberto no estudo da modelização e controlo do esquentador.

### 8.2.2 Dados para a Identificação

Na fase de aquisição de dados, tal como foi referido no capítulo 2, uma das preocupações é a definição correcta dos sinais de entrada a aplicar no sistema. Estes sinais devem possuir um espectro de frequências alargado e uma gama de amplitudes abrangentes. O procedimento habitual consiste em utilizar, como sinais excitadores da entrada sinais do tipo aleatório ou sinais com variações semelhantes às registadas durante a operação dos sistemas.

Durante a operação do esquentador os dois sinais de entrada (fluxo de gás e o caudal da água) apresentam diferentes tipos de variações. O sinal fluxo de gás varia em toda a gama de amplitudes e num espectro de frequências alargado; o sinal caudal da água varia, normalmente, sob a forma de uma “onda rectangular”.

Os sinais de excitação aplicados no esquentador foram definidos de forma a respeitar as variações típicas de operação referidas no parágrafo anterior, sendo a aquisição dos dados do sistema efectuada em malha aberta.

Os dados de entrada e saída foram ainda filtrados de forma a remover o ruído de medida de alta frequência sem alterar a informação real da dinâmica dos mesmos. Esta filtragem melhora a qualidade dos modelos identificados [Lju91].

Os dados foram divididos em dois grupos: treino e teste. Os dados de treino serão utilizados para o treino dos modelos e os dados de teste serão utilizados na fase de validação e verificação de capacidade de generalização dos modelos. É suposto que estes dois grupos de sinais sejam semelhantes, isto é, tenham formas e gamas de variação idênticas.

O conjunto de dados de treino apresentado na figura 8.1 e designado por  $\mathbf{D}^{ND}$  é definido pela equação 8.2:

$$\mathbf{D}^{ND} = \{f_g(k); c_a(k); \Delta t(k)\}, k = 1..ND \quad (8.2)$$

em que  $ND$  é o número de elementos de cada vector.

O conjunto de dados de teste apresentado na figura 8.2 e designado por  $\mathbf{H}^{NH}$  é definido pela equação 8.3:

$$\mathbf{H}^{NH} = \{f_g(k); c_a(k); \Delta t(k)\}, k = 1..NH \quad (8.3)$$

em que  $NH$  é o número de elementos de cada vector.

Foi efectuada ainda a recolha de mais dois conjuntos de dados estáticos e dinâmicos. No primeiro conjunto de dados, o caudal é fixo e provocaram-se variações em degrau no fluxo de gás. No segundo conjunto, o fluxo de gás é fixo e provocaram-se variações em degrau no caudal de água.

Estes dois conjuntos de dados destinam-se à visualização dos tipos de não linearidades existentes no esquentador e ainda à identificação dos modelos do tipo Hammerstein.

O conjunto de dados com caudal fixo e com variações em degrau do fluxo de gás apresentado na figura 8.3 e designado por  $\mathbf{E}^{NE}$  é definido pela equação 8.4:

$$\mathbf{E}^{NE} = \{f_g(k); c_a(k); \Delta t(k)\}, k = 1..NE \quad (8.4)$$

em que  $NE$  é o número de elementos de cada vector.

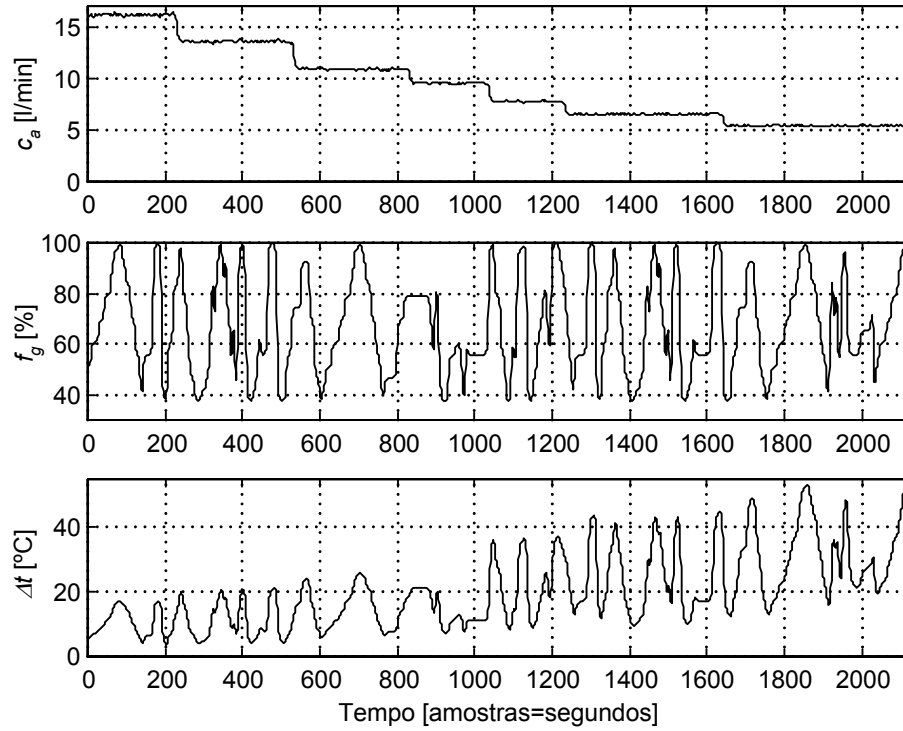


Figura 8.1: Dados de treino. De cima para baixo, gráfico 1 - caudal de água, gráfico 2 - fluxo de gás e gráfico 3 - aumento da temperatura.

O conjunto de dados com fluxo de gás fixo e com variações em degrau do caudal de água apresentado na figura 8.4 e designado por  $\mathbf{G}^{NG}$  é definido pela equação 8.5:

$$\mathbf{G}^{NG} = \{f_g(k); c_a(k); \Delta t(k)\}, k = 1..NG \quad (8.5)$$

em que  $NG$  é o número de elementos de cada vector.

O conjunto de dados  $\mathbf{G}^{NG}$  não foi utilizado por nenhum método de modelização apresentado neste trabalho. No entanto, é apresentado uma vez que, em conjunto com  $\mathbf{E}^{NE}$ , permite a observação das não linearidades do esquentador relativamente às duas variáveis de entrada fluxo de gás e caudal de água.

Durante a aquisição dos vários conjuntos de dados a temperatura da água fria permaneceu (praticamente) constante e igual a  $17,2^\circ C$ .

### 8.3 Modelo do Esquentador

A simplificação das entradas e saídas do esquentador, representada na figura 8.5 e apresentada no capítulo 7, permite descrever o comportamento do sistema, utilizando apenas duas variáveis de entrada: fluxo de gás ( $f_g(k)$ ) e caudal de água ( $c_a(k)$ ) e uma variável de saída: aumento da temperatura ( $\Delta t(k) = t_{aq}(k) - t_{af}(k)$ ).

Tendo por base esta simplificação e o estudo efectuado sobre os tempos mortos existentes no sistema, apresentado no capítulo 7, o modelo dinâmico global do esquentador

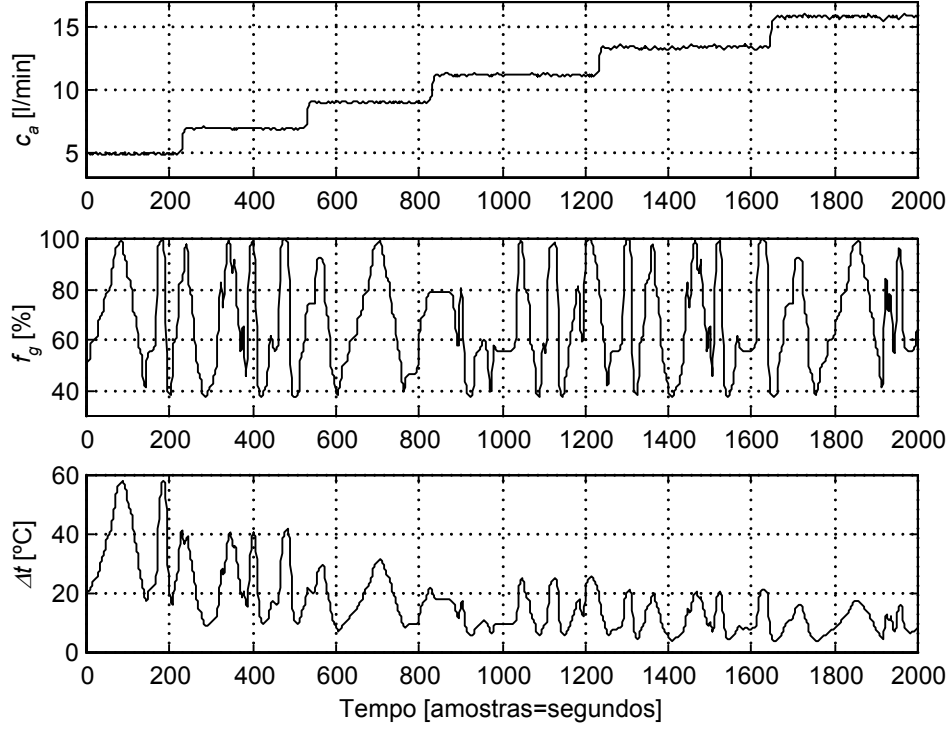


Figura 8.2: Dados de teste. De cima para baixo, gráfico 1 - caudal de água, gráfico 2 - fluxo de gás e gráfico 3 - aumento da temperatura.

é representado por dois blocos ligados em série: um bloco com o modelo dos tempos mortos e outro bloco com o modelo sem tempo morto.

### 8.3.1 Modelo dos Tempos Mortos

O tempo morto discreto entre a entrada e a saída de um dado sistema,  $d(k)$ , é calculado com base na equação 8.6:

$$d(k) = \left\lceil \frac{\tau_m(t)}{h} \right\rceil [h], k = t/h \quad (8.6)$$

em que  $\tau_m(t)$  é o tempo morto contínuo,  $h$  é o período de amostragem,  $k = t/h$  e  $\lceil x \rceil$  calcula o primeiro número inteiro maior ou igual a  $x$ . Na prática, o tempo morto de um dado sistema nunca é igual a um múltiplo inteiro exacto do intervalo de amostragem. Deste modo, o tempo morto de um dado sistema é constituído por duas componentes: uma inteira e outra fraccionária. A parte inteira está definida na equação 8.6 e a parte fraccionária, designada por  $m$ , está definida na equação 8.7:

$$\tau_m(k) = \left\lceil \frac{\tau_m(k)}{h} \right\rceil h - mh \quad [s] \quad (8.7)$$

em que  $0 \leq m < 1$  [Mot97].

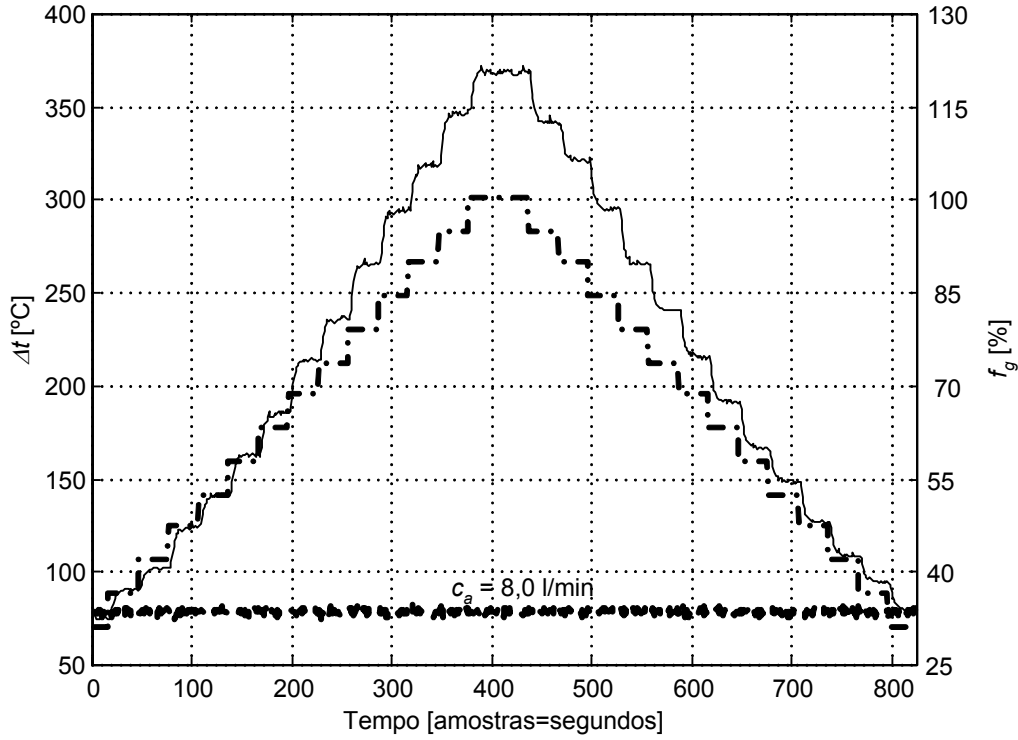


Figura 8.3: Dados de treino estático e dinâmico para um caudal fixo de 8 l/min (linha contínua - aumento da temperatura, linha traço ponto - fluxo de gás e linha tracejada - caudal de água).

Como se verifica no capítulo 7, o esquentador possui dois tempos mortos variáveis,  $\tau_{m1}(t)$  e  $\tau_{m2}(t)$ , definidos entre as duas variáveis de entrada fluxo de gás e caudal de água, e a variável de saída aumento da temperatura.

O tempo morto entre o fluxo de gás e o aumento da temperatura, apresentado no capítulo 7, é definido pela equação 8.8:

$$\tau_{m1}(t) = C_1 + \pi r^2 l \frac{C_2}{c_a(t)} = 2,50 + \pi r^2 l \frac{60000}{c_a(t)} \quad [\text{s}] \quad (8.8)$$

Substituindo  $C_1 = 2,5$  s (tempo morto constante entre a temperatura do permutador e o fluxo de gás),  $C_2 = 60000$  (factor de correcção de escala),  $r = 0,0078$  m (raio do tubo do permutador) e  $l = 0,29$  m (comprimento do tubo que a água percorre até atingir o sensor de água quente), a aproximação discreta do tempo morto variável  $d_1(k)$  é dada pela equação 8.9:

$$d_1(k) = \left\lceil \frac{2,50 + \frac{3,326}{c_a(k)}}{h} \right\rceil \quad [h] \quad (8.9)$$

Para os caudais admissíveis ( $[3,5; 14]$  l/min), o tempo morto variável  $d_1(k)$  é uma função estática que depende do caudal de água  $s_1(c_a(k))$  como definido na equação 8.9 e apresentado de uma forma explícita na equação 8.10:

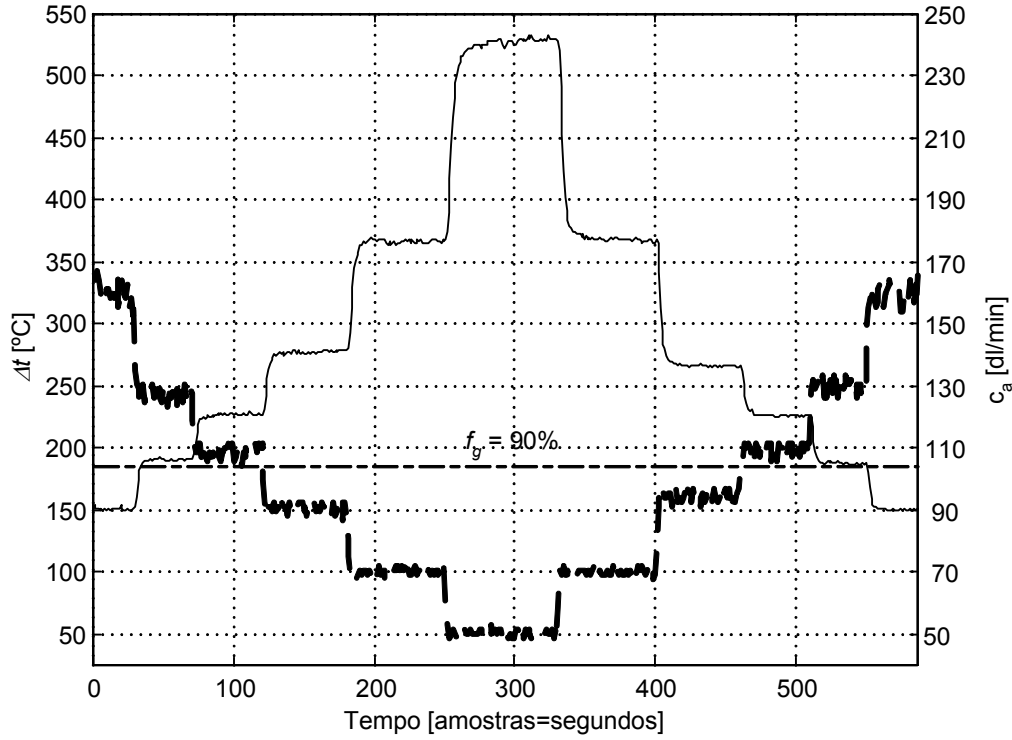


Figura 8.4: Dados de treinamento estático e dinâmico para um fluxo de gás de 90% (linha contínua - aumento da temperatura, linha traço ponto - fluxo de gás e linha tracejada - caudal de água).

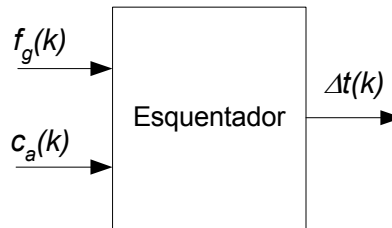


Figura 8.5: Entradas e saídas mensuráveis do esquentador.

$$d_1(k) = s_1(c_a(k)) = \begin{cases} 3; & 6,5 \leq c_a(k) < 14,0 \text{ l/min} \\ 4; & 3,5 \leq c_a(k) < 6,5 \text{ l/min} \end{cases} \quad [h] \quad (8.10)$$

O tempo morto  $\tau_{m1}(t)$ , além da parte inteira  $d_1(k)$ , possui uma parte fraccionária que não vai ser considerada no âmbito dos modelos desenvolvidos neste trabalho.

Além do tempo morto existente entre o fluxo de gás e o aumento da temperatura, o esquentador possui ainda um outro tempo morto entre o caudal de água e o aumento da temperatura. Este último é definido pela equação 8.11:

$$\tau_{m2}(t) = \pi r^2 l \frac{C_2}{c_a(t)} = \pi r l \frac{60000}{c_a(t)} \quad [\text{s}] \quad (8.11)$$

Substituindo  $C_2 = 60000$  (factor de correcção de escala),  $r = 0,0078$  m (raio do tubo do permutador) e  $l = 0,29$  m (comprimento do tubo que a água percorre até atingir o sensor de água quente), a aproximação discreta do tempo morto variável  $d_2(k)$ , na iteração  $k$ , é dada pela equação 8.12:

$$d_2(k) = \left\lceil \frac{\frac{3,326}{c_a(k)}}{h} \right\rceil \quad [h] \quad (8.12)$$

Para a gama de caudais admissíveis, o tempo morto variável discreto  $d_2(k)$  é uma função estática que só depende do caudal de água  $s_2(c_a(k))$  como definido na equação 8.12 e apresentado de uma forma explícita na equação 8.13:

$$d_2(k) = s_2(c_a(k)) = \begin{cases} 1; & 3,5 \leq c_a(k) < 14,0 \end{cases} \quad 1/\text{min} \quad [h] \quad (8.13)$$

Neste caso, o tempo morto relativamente ao caudal da água é praticamente constante, daí ser considerado igual a um período de amostragem  $d_2(k) = 1$ .

O tempo morto  $\tau_{m2}(t)$ , além da parte inteira  $d_2(k)$ , possui uma parte fraccionária que não vai ser considerada no âmbito dos modelos desenvolvidos neste trabalho.

A figura 8.6 apresenta, de forma gráfica, os dois tempos mortos ( $\tau_{m1}(t)$  e  $\tau_{m2}(t)$  valores teóricos, a linhas contínuas, e  $d_1(k)$  e  $d_2(k)$  respectivos valores discretos a  $^\circ$  e  $^*$ , respectivamente).

O diagrama de blocos ilustrado na figura 8.7 representa o modelo dos tempos mortos do esquentador relativos às duas variáveis de entradas.

### 8.3.2 Modelo do Esquentador sem Tempo Morto

O aumento da temperatura da água na iteração  $k+1$  é uma função matemática  $m(\cdot)$  que depende do fluxo de gás na iteração  $k - d_1(k)$ , do caudal de água na iteração  $k - d_2(k)$  e do aumento da temperatura da água na iteração  $k$ , como definido na equação 8.14 e representado no diagrama de blocos da figura 8.8:

$$\widehat{\Delta t}(k+1) = m(\widehat{\Delta t}(k), f_g(k - d_1(k)), c_a(k - d_2(k))) \quad (8.14)$$

em que  $k$  representa a amostra no instante de tempo  $t = kh$ ,  $h$  é o período de amostragem e  $d_1(k)$  e  $d_2(k)$  são os tempos mortos relativos às duas entradas.

### 8.3.3 Modelo Global do Esquentador

O diagrama de blocos do modelo global do esquentador, apresentado na figura 8.9, é constituído pela junção dos dois blocos atrás apresentados, o bloco com o modelo dos tempos mortos e o bloco com o modelo sem tempo morto.

Daqui em diante sempre que se fizer referência ao modelo do esquentador está a fazer-se referência ao modelo do esquentador sem tempo morto.



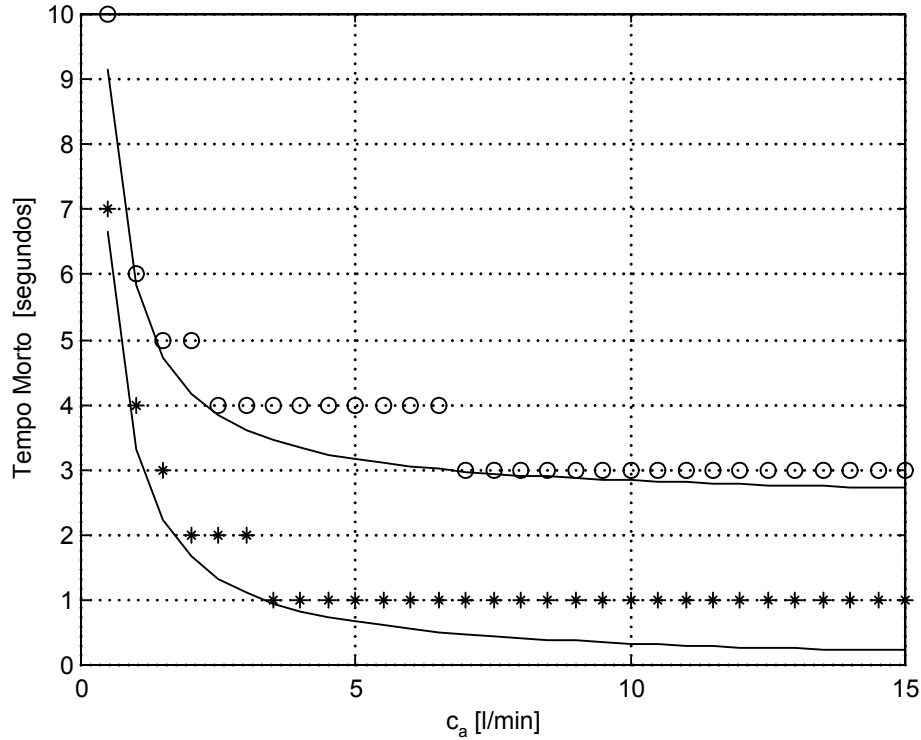


Figura 8.6: Tempos mortos relativos ao fluxo de gás e ao caudal de água ( $\tau_{m1}(t)$  e  $\tau_{m2}(t)$  valores teóricos contínuos - linhas contínuas e  $d_1(k)$  e  $d_2(k)$  valores discretos -  $\circ$  e  $*$ , respectivamente).

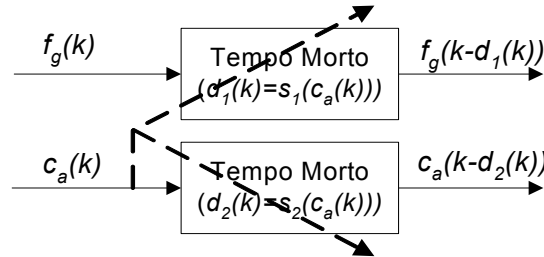


Figura 8.7: Modelo dos tempos mortos variáveis do esquentador.

## 8.4 Estruturas de Treino

### 8.4.1 Identificação do Modelo Directo

O diagrama de blocos da estrutura utilizada para a identificação do modelo directo do esquentador está ilustrada na figura 8.10. A saída do modelo é designada por  $\widehat{\Delta t}(k+1)$  e a saída real por  $\Delta t(k+1)$ .

A estimação da saída do modelo  $\widehat{\Delta t}(k+1)$  é função de  $m(\cdot)$  e dos regressores  $\Delta t(k)$ ,  $f_g(k - d_1(k))$  e  $c_a(k - d_2(k))$ , como definido na equação 8.15:

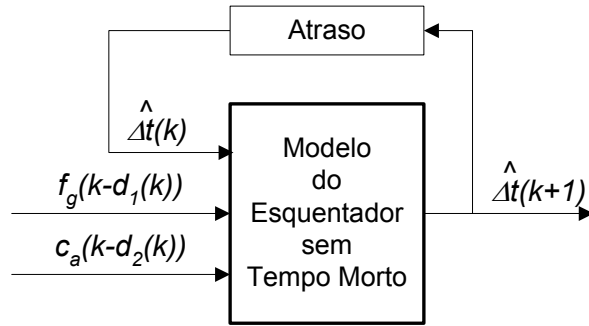


Figura 8.8: Modelo do esquentador sem tempo morto.

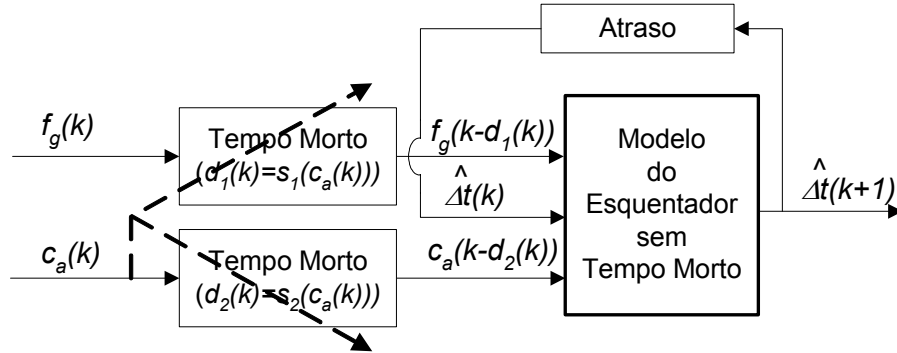


Figura 8.9: Modelo directo global do esquentador.

$$\widehat{\Delta t}(k+1) = m(\Delta t(k), f_g(k-d_1(k)), c_a(k-d_2(k))) \quad (8.15)$$

Com base numa função de custo a definir baseada no erro  $e(k+1) = \Delta t(k+1) - \widehat{\Delta t}(k+1)$ , no conjunto de dados de treino  $\mathbf{D}^{ND}$  e no método de identificação escolhido, os parâmetros do modelo directo do esquentador são identificados.

### 8.4.2 Identificação do Modelo Inverso

O diagrama de blocos da estrutura utilizada para a identificação do modelo inverso está ilustrada na figura 8.11. A saída do modelo inverso é designada por  $\widehat{f}_g(k-d_1(k))$  e a variável real correspondente por  $f_g(k-d_1(k))$ .

A estimação da saída do modelo  $\widehat{f}_g(k-d_1(k))$  é função de  $m^{-1}(\cdot)$  e dos regressores  $\Delta t(k+1)$ ,  $\Delta t(k)$  e  $c_a(k-d_2(k))$ , como definido na equação 8.16:

$$\widehat{f}_g(k-d_1(k)) = m^{-1}(\Delta t(k+1), \Delta t(k), c_a(k-d_2(k))) \quad (8.16)$$

Com base numa função de custo a definir baseada no erro  $e(k-d_1(k)) = f_g(k-d_1(k)) - \widehat{f}_g(k-d_1(k))$ , no conjunto de dados de treino  $\mathbf{D}^{ND}$  e no método de identificação escolhido, os parâmetros do modelo inverso são identificados.

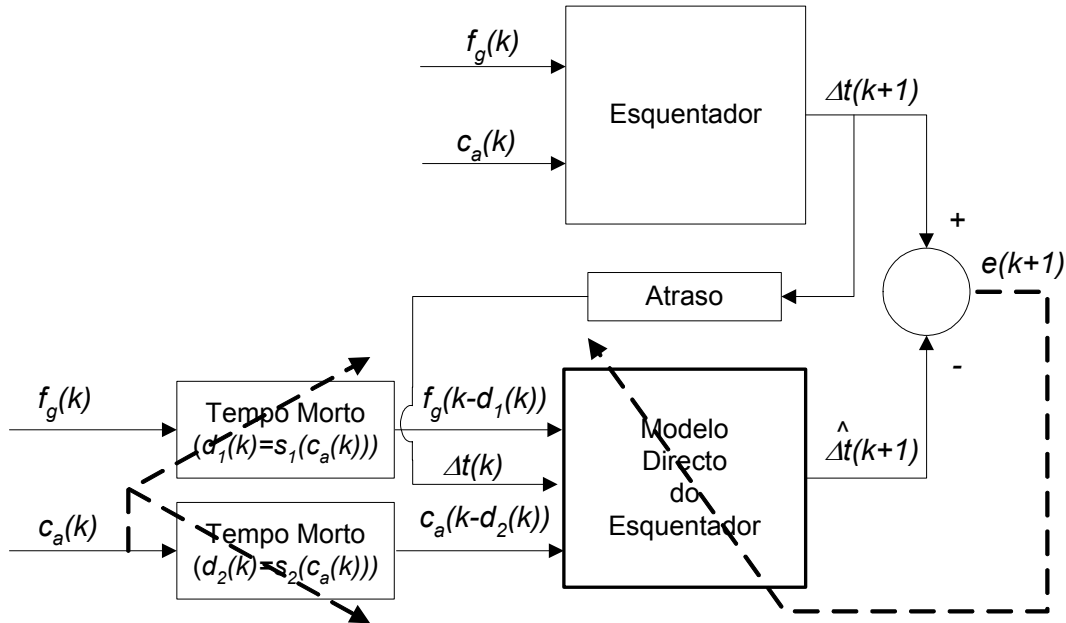


Figura 8.10: Estrutura de treino do modelo directo global do esquentador.

O modelo inverso identificado desta forma é conhecido por modelo inverso genérico [HSZG92]. O facto deste modelo ser identificado com base no erro entre os dois fluxos de gás e não com base no erro entre os dois aumentos de temperatura (variáveis de saída) pode trazer problemas na identificação do modelo inverso, principalmente quando se estão a utilizar métodos de modelização do tipo caixa negra ou cinzenta escura. Estes problemas podem surgir por dois motivos: primeiro, porque não existe garantia de que os sinais de entrada utilizados sejam semelhantes aos sinais de funcionamento em malha fechada; segundo, porque o sistema a modelizar pode não ser biunívoco. Assim, a identificação modelo inverso efectuada deste modo pode não ser única [MO95].

Para evitar os problemas apontados, existe uma outra estrutura de identificação do modelo inverso conhecida por estrutura de modelo inverso especializado [HSZG92], em que o modelo inverso é colocado como controlador em série com o sistema, sendo neste aplicados sinais de referência idênticos aos utilizados no controlo do sistema. O erro entre estes sinais e a saída do sistema é usado na identificação dos parâmetros do modelo inverso com base na informação do modelo directo.

Neste trabalho, o modelo inverso é definido e identificado pelo método genérico, uma vez que o modelo inverso do esquentador existe sempre e é único (ver nas secções seguintes). O modelo inverso do esquentador é definido sempre que possível através da inversão matemática do modelo directo. Este procedimento, além de permitir a definição do modelo inverso “correcto”, garante ainda o emparelhamento dos modelos directo e inverso.

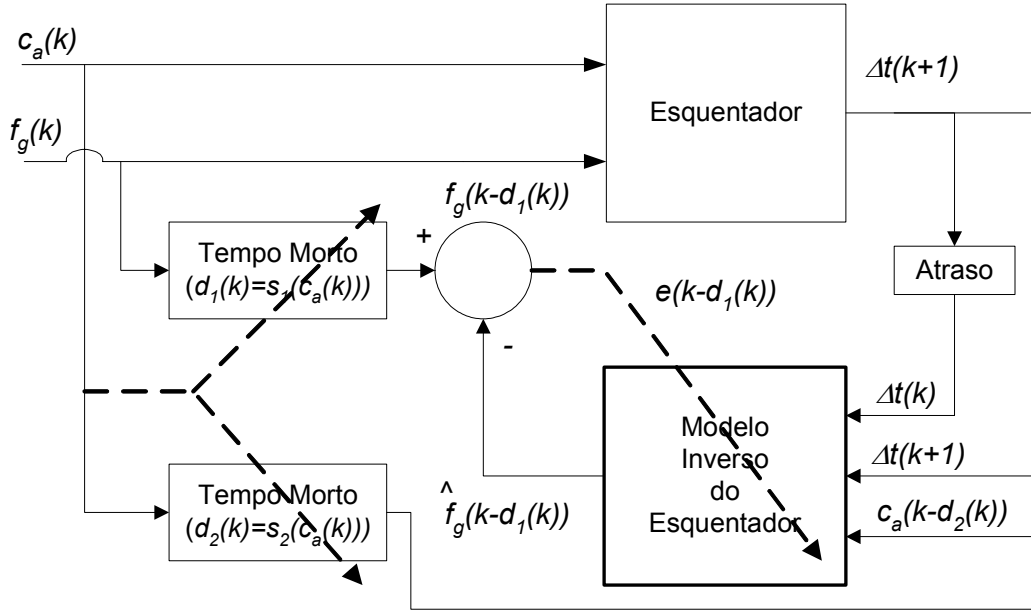


Figura 8.11: Estrutura de treino do modelo inverso global do esquentador.

## 8.5 Modelização e Identificação Linear

O primeiro método de modelização utilizado para a construção de um modelo do esquentador é a modelização linear.

### 8.5.1 Modelo Linear

#### Modelo Linear Directo

Com os regressores definidos na equação 8.15, a estimação da saída  $\widehat{\Delta t}(k+1)$  do modelo linear de 1ª ordem (capítulo 7) é dada pela equação 8.17:

$$\widehat{\Delta t}(k+1) = a_1 \Delta t(k) + b_1 f_g(k - d_1(k)) + c_1 c_a(k - d_2(k)) \quad (8.17)$$

em que  $\theta = [a_1, b_1, c_1]^T$  é o vector dos parâmetros lineares a identificar.

#### Modelo Linear Inverso

O modelo inverso pode ser definido através da inversão matemática do modelo directo, sendo a saída deste  $\widehat{f}_g(k - d_1(k))$ , dada pela inversão da equação 8.17, resultando na equação 8.18:

$$\widehat{f}_g(k - d_1(k)) = \frac{-\Delta t(k+1) + a_1 \Delta t(k) + c_1 c_a(k - d_2(k))}{-b_1} \quad (8.18)$$

### 8.5.2 Identificação do Modelo Linear

#### Modelo Directo

A identificação do modelo linear é efectuada minimizando a função de custo  $J(\boldsymbol{\theta}, \mathbf{D}^{ND})$  definida na equação 8.19:

$$J(\boldsymbol{\theta}, \mathbf{D}^{ND}) = \frac{1}{ND} \sum_{k=1}^{ND} e(k)^2 \quad (8.19)$$

em que  $e(k)$  é a diferença entre a saída real e a saída estimada e  $\mathbf{D}^{ND}$  os dados de treino. Recorrendo à estrutura de identificação apresentada na figura 8.10 e ao algoritmo de identificação de mínimos quadrados, os parâmetros  $\boldsymbol{\theta} = [a_1, b_1, c_1]^T$  do modelo linear foram identificados.

#### Modelo Inverso

Como se apresentou atrás, o modelo inverso não necessita de ser identificado utilizando os dados de treino, uma vez que é possível inverter matematicamente o modelo directo atrás identificado, desde que os três parâmetros sejam reais finitos e  $b_1$  seja diferente de zero.

### 8.5.3 Resultados

A validação dos modelos é efectuada utilizando a função raiz da média do quadrado dos erros (mais vulgarmente conhecida por *Root Mean Square Error* - *RMSE*) definida na equação 8.20:

$$V(\boldsymbol{\theta}, \mathbf{H}^{NH}) = \sqrt{\frac{1}{NH} \sum_{k=1}^{NH} e(k)^2} \quad (8.20)$$

em que  $e(k)$  é o erro entre o valor de saída real e estimada e  $\mathbf{H}^{NH}$  são os dados de teste.

A função  $V(\boldsymbol{\theta}, \mathbf{H}^{NH})$  calcula a média do quadrado dos erros (mais vulgarmente conhecida por *Mean Square Error* - *MSE*) em função dos parâmetros  $\boldsymbol{\theta}$  e dos dados de teste  $\mathbf{H}^{NH}$ .

A tabela 8.1 apresenta os valores de  $\sqrt{J(\boldsymbol{\theta}, \mathbf{D}^{ND})}$  e  $V(\boldsymbol{\theta}, \mathbf{H}^{NH})$  obtidos, respectivamente, para os modelos directo e inverso.

Função de Custo	Modelo	
	Directo [ $^{\circ}C$ ]	Inverso [%]
Treino - $\sqrt{J(\boldsymbol{\theta}, \mathbf{D}^{ND})}$	4,50	11,00
Teste - $V(\boldsymbol{\theta}, \mathbf{H}^{NH})$	4,90	11,50

Tabela 8.1: MSEs obtidos com o modelo linear.

### Modelo Directo

Na figura 8.12 estão apresentados os resultados obtidos pelo modelo linear directo. De cima para baixo, o gráfico 1 apresenta o aumento da temperatura estimada e real (linha tracejada e contínua, respectivamente) e o gráfico 2 apresenta o erro entre os dois sinais referidos, utilizando os dados de treino.

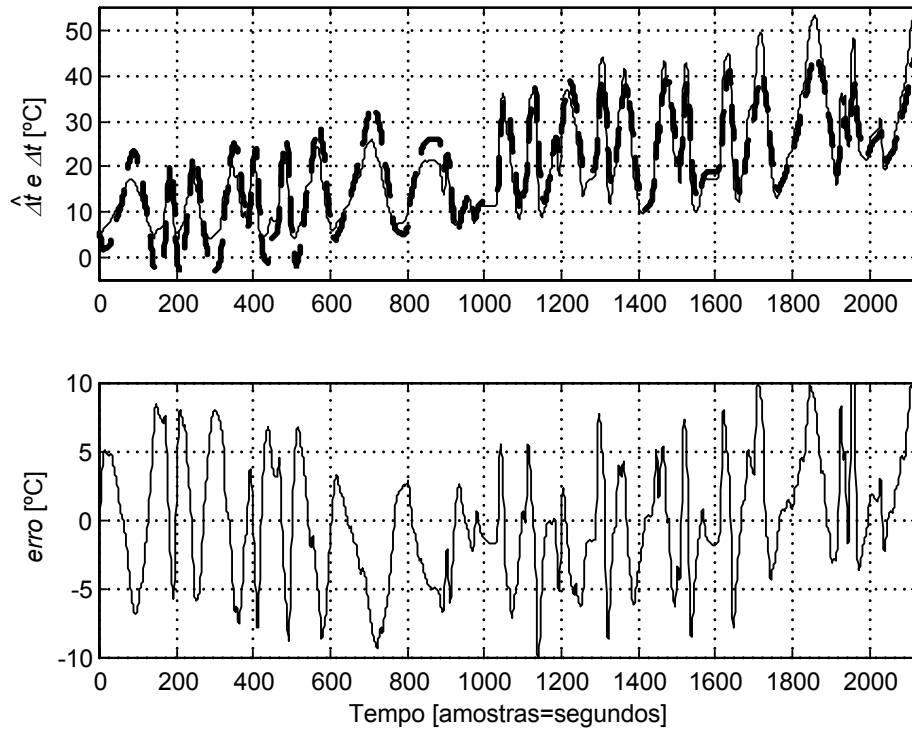


Figura 8.12: Resultados do modelo directo linear. De cima para baixo, gráfico 1 - aumento da temperatura estimada e real (linha tracejada e contínua, respectivamente) e gráfico 2 - sinal de erro entre os dois sinais anteriores, em ordem ao tempo.

### Modelo Inverso

Os resultados obtidos pelo modelo linear inverso estão apresentados na figura 8.13. De cima para baixo, o gráfico 1 apresenta o fluxo de gás estimado e real (linha tracejada e contínua, respectivamente) e o gráfico 2 apresenta o erro entre os dois sinais referidos, utilizando os dados de treino.

#### 8.5.4 Comentários

Pelas análises das figuras 8.12 e 8.13 pode observar-se que os sinais reais e estimados apresentam diferenças significativas.

Apesar dos modelos lineares do esquentador apresentarem estruturas e algoritmos de aprendizagem muito simples, eles possuem; no entanto, elevados erros de modelização, como se pode ver pela tabela 8.1. O insucesso da aplicação deste método de

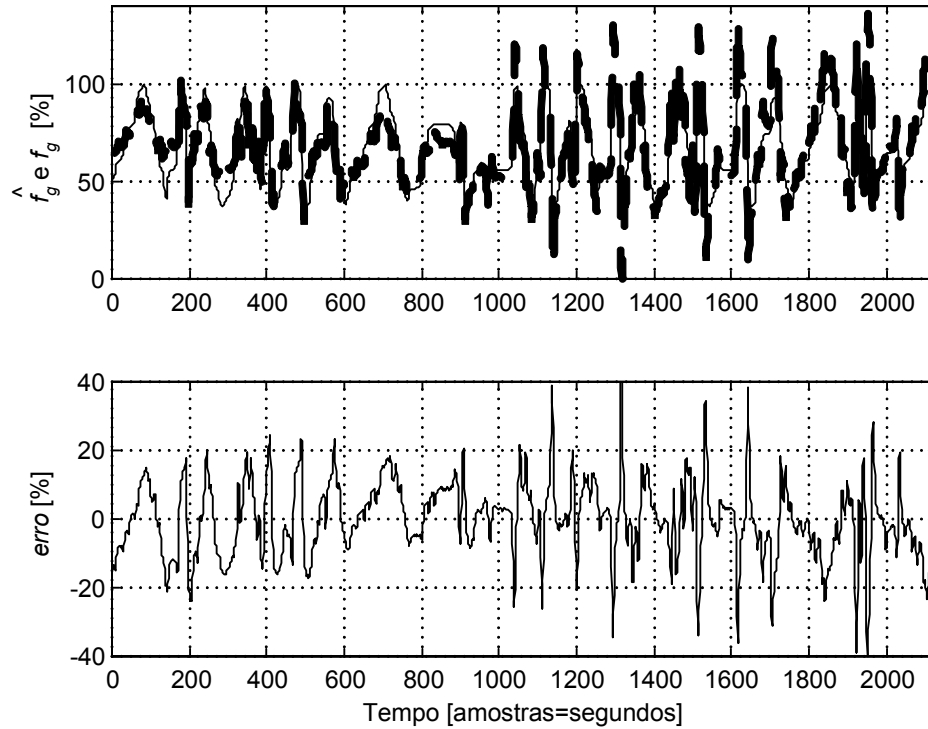


Figura 8.13: Resultados do modelo inverso linear. De cima para baixo, gráfico 1 - fluxo de gás estimado e real (linha tracejada e contínua, respectivamente) e gráfico 2 - sinal de erro entre os dois sinais anteriores, em ordem ao tempo.

modelização deve-se ao facto do esquentador se tratar de um sistema claramente não linear.

## 8.6 Modelização e Identificação Neuro-Difusa

O primeiro método de modelização não linear utilizado é a modelização neuro-difusa apresentada no capítulo 3. Os modelos são construídos com base num conjunto de regras linguísticas que surgem de conhecimentos heurísticos e experimentais, construindo modelos com um determinado nível de interpretabilidade. Contudo, a identificação dos parâmetros internos destes modelos é efectuada com base num conjunto de dados de entrada e saída podendo este procedimento reduzir o seu nível de interpretabilidade.

O desenvolvimento de modelos neuro-difusos não exige grandes conhecimentos *a priori* sobre o sistema e o tempo dispendido no seu desenvolvimento é pequeno. Em contrapartida, a interpretabilidade destes modelos é relativamente baixa e a complexidade matemática é relativamente elevada. Estes modelos são classificados como modelos do tipo caixa cinzenta escura.

Dada a abundância de ferramentas de *software* para a modelização neuro-difusa, foram estudadas e comparadas três ferramentas para a construção de modelos neuro-difusos do esquentador: ANFIS [Jan93], ANNBFIS [LC97] e NFIDA [BB99]. Estas

três ferramentas implementam modelos difusos do tipo Takagi-Sugeno de ordem um ou zero, residindo a principal diferença nos métodos de inicialização das respectivas estruturas (número, posição e forma dos conjuntos difusos e definição das regras). Os resultados desta comparação foram publicados em [VM03a]. Neles se concluiu que os modelos construídos apresentam resultados muito semelhantes no que respeita ao erro de modelização e à complexidade matemática.

Por este facto, a modelização neuro-difusa do esquentador apresentada nesta dissertação é efectuada utilizando apenas a ferramenta ANFIS. Esta ferramenta disponibiliza vários tipos de sistemas neuro-difusos, várias formas de inicialização das respectivas estruturas e vários algoritmos de treino e paragem, permitindo ao utilizador explorar as várias capacidades deste tipo de modelos de uma forma fácil e expedita.

### 8.6.1 Modelo Neuro-Difuso

#### Modelo Neuro-Difuso Directo

O modelo neuro-difuso é definido por um conjunto de quatro regras difusas do tipo Takagi-Sugeno de ordem um e os regressores apresentados na equação 8.15. As regras difusas estão definidas na equação 8.21:

$$\begin{aligned} & - \text{Se } \Delta t(k) \text{ é } A_{il} \text{ e } f_g(k - d_1(k)) \text{ é } A_{il} \text{ e } c_a(k - d_2(k)) \text{ é } A_{il}, \\ & \text{Então} \\ & \Delta t_i(k + 1) = a_i \Delta t(k) + b_i f_g(k - d_1(k)) + c_i c_a(k - d_2(k)) + m_i; \end{aligned} \quad \begin{array}{l} \text{com } i = 1..4 \\ \text{e } l = 1..3 \end{array} \quad (8.21)$$

em que  $i$  indica o número da regra e  $l$  a respectiva entrada. As funções de pertença dos antecedentes são funções Gaussianas definidas pela equação 8.22:

$$\mu_{A_{il}}(x_l; c_{il}; \sigma_{il}) = e^{\left(-\frac{(x_l - c_{il})^2}{2\sigma_{il}^2}\right)}, \quad i = 1..4 \text{ e } l = 1..3 \quad (8.22)$$

onde  $c_{il}$  e  $\sigma_{il}$  são os centros e os desvios padrão das várias funções de pertença relativas a cada entrada  $l$  e respeitantes a cada regra  $i$ .

As funções de pertença dos consequentes são funções lineares relativamente às entradas do modelo definidas pela equação 8.23:

$$\Delta t_i(k + 1) = a_i \Delta t(k) + b_i f_g(k - d_1(k)) + c_i c_a(k - d_2(k)) + m_i \quad (8.23)$$

O grau de pertença global de cada regra  $\beta_i$  é dado pelo produto algébrico (operador  $\prod$ ) dos graus de pertença dos conjuntos difusos correspondentes às várias entradas da referida regra, como definido na equação 8.24:

$$\beta_i = \prod_{l=1}^3 \mu_{A_{il}}(x_l; c_{il}; \sigma_{il}), \quad i = 1..4 \quad (8.24)$$

A operação normalização dos pesos (operador  $\mathbf{N}$ ) é obtida conforme definido na equação 8.25:



$$\gamma_i = \frac{\beta_i}{\sum_{i=1}^4 \beta_i} = \frac{\prod_{l=1}^3 e^{-\frac{(x_l - c_{il})^2}{2\sigma_{il}^2}}}{\sum_{i=1}^4 \left( \prod_{l=1}^3 e^{-\frac{(x_l - c_{il})^2}{2\sigma_{il}^2}} \right)}, \quad i = 1..4 \quad (8.25)$$

Desta forma, aplicando o operador produto algébrico aos consequentes, a saída de cada regra é o produto algébrico entre o grau de pertença normalizado  $\gamma_i$  e a função linear da respectiva regra, como definido na equação 8.26:

$$\gamma_i \Delta t_i(k+1) = \gamma_i (a_i \Delta t(k) + b_i f_g(k - d_1(k)) + c_i c_a(k - d_2(k)) + m_i), \quad i = 1..4 \quad (8.26)$$

Finalmente, a saída global do modelo neuro-difuso  $\widehat{\Delta t}(k+1)$  é dada pela soma linear das saídas de cada regra  $\Delta t_i(k+1)$ , conforme o apresentado na equação 8.27:

$$\begin{aligned} \widehat{\Delta t}(k+1) &= \sum_{i=1}^4 \gamma_i (a_i \Delta t(k) + b_i f_g(k - d_1(k)) + c_i c_a(k - d_2(k)) + m_i) \\ &= \sum_{i=1}^4 \left( \frac{\prod_{l=1}^3 e^{-\frac{(x_l - c_{il})^2}{2\sigma_{il}^2}}}{\sum_{i=1}^4 \left( \prod_{l=1}^3 e^{-\frac{(x_l - c_{il})^2}{2\sigma_{il}^2}} \right)} (a_i \Delta t(k) + b_i f_g(k - d_1(k)) + c_i c_a(k - d_2(k)) + m_i) \right) \end{aligned} \quad (8.27)$$

O modelo neuro-difuso directo é constituído por quatro regras cujos respectivos consequentes são definidos por quatro parâmetros lineares, resultando num total de dezasseis parâmetros lineares ( $a_i$ ,  $b_i$ ,  $c_i$  e  $m_i$  com  $i = 1..4$ ). Quanto aos parâmetros não lineares, o modelo possui três entradas representadas através de quatro conjuntos difusos diferentes definidos por dois parâmetros não lineares, resultando num total de vinte e quatro parâmetros não lineares ( $c_{il}$  e  $\sigma_{il}$  com  $i = 1..4$  e  $l = 1..3$ ). Deste modo, este modelo possui um total de quarenta parâmetros.

### Modelo Neuro-Difuso Inverso

Na impossibilidade de inverter a função do modelo neuro-difuso directo (equação 8.27), o modelo inverso é construído com uma estrutura idêntica à utilizada no modelo neuro-difuso directo (igual número de regras e parâmetros), utilizando os regressores apresentados na equação 8.16.

Este modelo possui igualmente um total de quarenta parâmetros diferentes do modelo directo.

## 8.6.2 Identificação do Modelo Neuro-Difuso

### Identificação do Modelo Directo

A definição do número e posição das funções de pertença do tipo Gaussiano e a definição e o número das regras é efectuada utilizando o algoritmo de agrupamento de dados subtrativo [Chi94] apresentado no capítulo 3. Neste caso, este algoritmo agrupa os dados de treino em quatro grupos criando quatro regras. Os centros dos quatro grupos são os centros das funções de pertença e os desvios padrão indicam o grau de sobreposição entre os grupos.

O modelo neuro-difuso é identificado com base na função de custo definida na equação 8.19 e recorrendo à estrutura de identificação apresentada na figura 8.10. O algoritmo de identificação dos parâmetros da estrutura utilizado é o algoritmo híbrido iterativo semelhante ao apresentado no capítulo 3. Cada iteração deste algoritmo é composta por duas partes: primeiro são identificados os parâmetros lineares pelo método dos mínimos quadrados; de seguida, são identificados os parâmetros não lineares pelo método do gradiente. Para reduzir a carga computacional deste algoritmo de aprendizagem, mantendo a sua velocidade de convergência, a primeira parte do algoritmo só é executada de dez em dez iterações. Para aumentar a convergência deste algoritmo, o factor de aprendizagem é alterado consoante a diminuição do erro de modelização [Mat01]. A paragem deste algoritmo de aprendizagem é efectuada quando a função de custo, relativamente aos dados de teste, deixa de diminuir.

### Identificação do Modelo Inverso

A identificação do modelo inverso é realizada da mesma forma que a identificação do modelo directo, com a excepção dos regressores utilizados (equação 8.16) e da estrutura de treino (figura 8.11). A estrutura e o número de parâmetros do modelo neuro-difuso inverso são iguais aos do modelo directo.

## 8.6.3 Resultados

A tabela 8.2 apresenta os valores de  $\sqrt{J(\boldsymbol{\theta}, \mathbf{D}^{ND})}$  e  $V(\boldsymbol{\theta}, \mathbf{H}^{NH})$  obtidos, respectivamente, para os modelos directo e inverso.

Função de Custo	Modelo	
	Directo [ $^{\circ}C$ ]	Inverso [%]
Treino - $\sqrt{J(\boldsymbol{\theta}, \mathbf{D}^{ND})}$	0,36	1,17
Teste - $V(\boldsymbol{\theta}, \mathbf{H}^{NH})$	0,59	1,28

Tabela 8.2: MSEs obtidos com o modelo neuro-difuso.

### Modelo Directo

Os resultados obtidos pelo modelo directo neuro-difuso estão apresentados na figura 8.14. De cima para baixo, o gráfico 1 apresenta o aumento da temperatura estimado e

real (linha tracejada e contínua, respectivamente) e o gráfico 2 apresenta o erro entre os dois sinais referidos, utilizando os dados de treino.

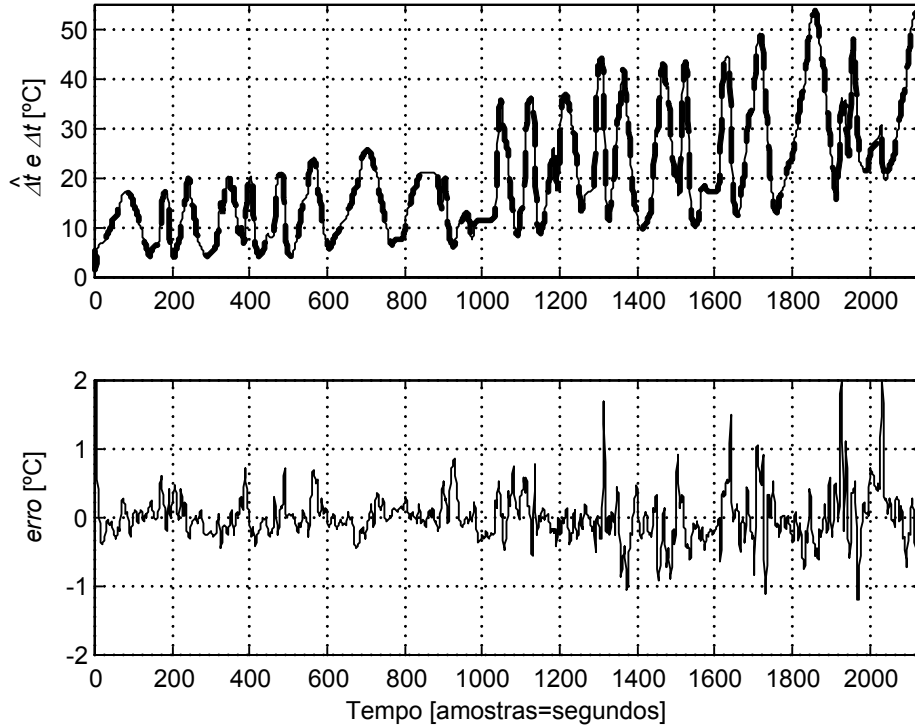


Figura 8.14: Resultados do modelo directo neuro-difuso. De cima para baixo, gráfico 1 - aumento da temperatura estimada e real (linha tracejada e contínua, respectivamente) e gráfico 2 - sinal de erro entre os dois sinais anteriores, em ordem ao tempo.

### Modelo Inverso

Na figura 8.15 estão apresentados os resultados obtidos pelo modelo inverso neuro-difuso. De cima para baixo, o gráfico 1 apresenta o fluxo de gás estimado e real (linha tracejada e contínua, respectivamente) e o gráfico 2 o erro entre os dois sinais referidos, utilizando os dados de treino.

#### 8.6.4 Comentários

Como se pode ver pelas figuras 8.14 e 8.15 os sinais reais e estimados pelo modelo estão praticamente sobrepostos, o que indicia uma boa modelização. Pode ainda comprovar-se pelos resultados apresentados na tabela 8.2, que os modelos neuro-difusos obtêm erros de modelização baixos. Contudo, esta solução dá origem a modelos complexos (elevado número de parâmetros) e com um nível de interpretabilidade relativamente baixo.

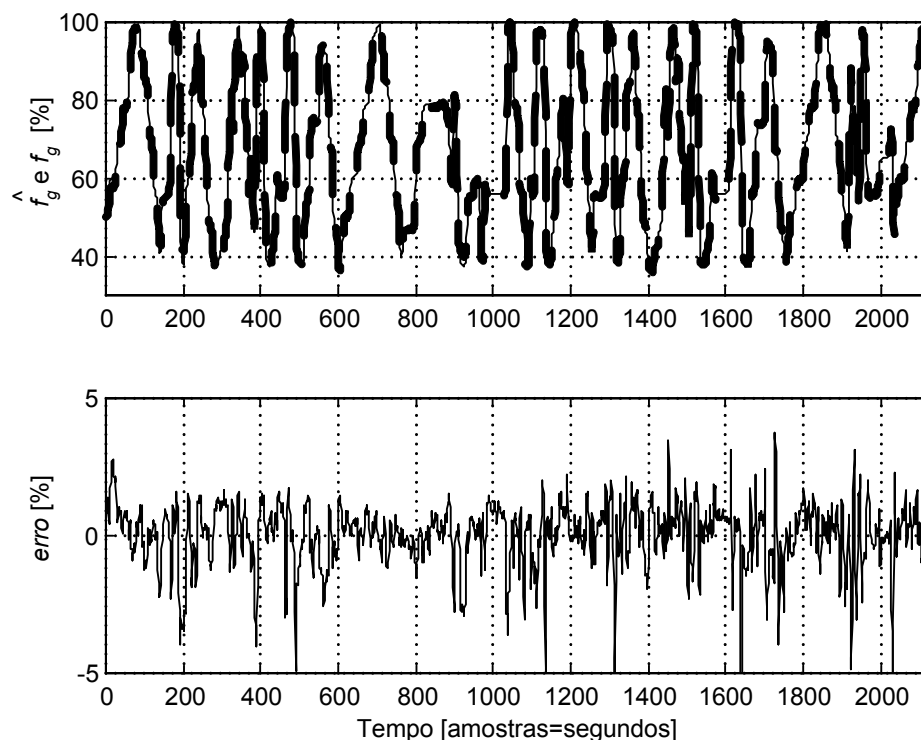


Figura 8.15: Resultados do modelo inverso neuro-difuso. De cima para baixo, gráfico 1 - fluxo de gás estimado e real (linha tracejada e contínua, respectivamente) e gráfico 2 - sinal de erro entre os dois sinais anteriores, em ordem ao tempo.

## 8.7 Modelização e Identificação Orientada por Blocos

Com o recurso a uma maior quantidade de conhecimentos adquiridos *a priori* sobre o funcionamento do esquentador verifica-se que, para um determinado caudal fixo, o esquentador pode ser modelizado através da ligação em série de uma função não linear estática  $f(\cdot)$ , do tipo quadrático, com um sistema linear dinâmico de 1ª ordem  $G(\cdot)$ . Os modelos representados desta forma são conhecidos por modelos orientados por blocos, mais especificamente por modelos Hammerstein.

Apesar dos modelos orientados por blocos não se basearem fortemente em conhecimentos físicos, o facto destes serem constituídos pela conjugação de dois blocos com algum significado físico dá-lhes um razoável nível de interpretabilidade (daí o seu enquadramento nos métodos de modelização cinzenta).

### 8.7.1 Modelo Hammerstein

Com base nos conhecimentos apresentados no capítulo 7 e nos dados estáticos e dinâmicos do esquentador, apresentados respectivamente nas figuras 8.3 e 8.1, verifica-se que, para a um determinado caudal fixo, o esquentador pode ser modelizado pela li-

gação em série de uma função não linear estática e de um modelo linear dinâmico de 1ª ordem, como ilustra o diagrama de blocos da figura 8.16.

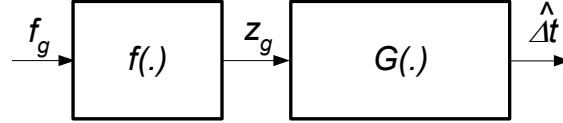


Figura 8.16: Modelo Hammerstein do esquentador.

Neste caso, o modelo do esquentador não possui o caudal de água como variável de entrada, mesmo sabendo que o modelo global do esquentador depende fortemente desta variável. Como a dependência é não linear e desconhecida, a opção é identificar em linha, recursivamente, os parâmetros do modelo linear de 1ª ordem, uma vez que a constante de tempo e o ganho do esquentador dependem fortemente da referida variável.

Os regressores utilizados para a construção do modelo Hammerstein directo são os apresentados na equação 8.28:

$$\widehat{\Delta t}(k+1) = m(\Delta t(k), f_g(k - d_1(k))) \quad (8.28)$$

Os regressores utilizados para a construção do modelo Hammerstein inverso são os apresentados na equação 8.29:

$$\widehat{f}_g(k - d_1(k)) = m^{-1}(\Delta t(k+1), \Delta t(k)) \quad (8.29)$$

A função não linear estática é identificada, fora de linha, utilizando duas funções diferentes: função polinomial e sistema neuro-difuso, dando origem a duas variantes de modelos Hammerstein: modelo Hammerstein adaptativo polinomial e modelo Hammerstein adaptativo neuro-difuso.

A modelização da função não linear estática utilizando sistemas neuro-difusos é mais geral, uma vez que estes permitem a modelização de qualquer função não linear. A modelização da referida função utilizando funções polinomiais é mais restritiva.

Uma vez que os parâmetros do modelo linear de 1ª ordem são adaptados recursivamente em linha, os modelos apresentados são, neste contexto, designados por modelos Hammerstein adaptativos.

### Modelo Hammerstein Adaptativo Polinomial

**Modelo Directo** O modelo Hammerstein adaptativo polinomial é constituído por uma função polinomial estática  $f(\cdot)$  em série com um modelo linear dinâmico. Como se pode ver pela figura 8.16, este modelo possui uma variável intermédia designada por  $z_g$ , que é a saída da função estática não linear e, simultaneamente, a entrada do modelo linear dinâmico (variável não mensurável directamente, mas relacionada com o aumento da temperatura em estado estacionário).

Baseado nos conhecimentos existentes sobre a parte estática do esquentador, apresentados no capítulo 7, verifica-se que uma função polinomial de ordem dois é suficiente para a modelização correcta de  $f(\cdot)$ :

$$\begin{aligned} z_g(k - d_1(k)) &= f(f_g(k - d_1(k))) \\ z_g(k - d_1(k)) &= a_0 + a_1 f_g(k - d_1(k)) + a_2 f_g(k - d_1(k))^2 \end{aligned} \quad (8.30)$$

O modelo linear dinâmico de 1<sup>a</sup> ordem utilizado é o modelo definido na equação 8.31:

$$\widehat{\Delta t}(k + 1) = p_1 \Delta t(k) + q_1 z_g(k - d_1(k)) \quad (8.31)$$

em que  $\theta = [p_1, q_1]^T$  é o vector dos parâmetros lineares.

A equação 8.32 define, deste modo, o modelo Hammerstein adaptativo polinomial:

$$\widehat{\Delta t}(k + 1) = p_1 \Delta t(k) + q_1 (a_0 + a_1 f_g(k - d_1(k)) + a_2 f_g(k - d_1(k))^2) \quad (8.32)$$

Os parâmetros  $p_1$  e  $q_1$  são parâmetros variáveis a identificar em linha, e  $a_0$ ,  $a_1$  e  $a_2$  são parâmetros constantes a identificar fora de linha.

**Modelo Inverso** O modelo inverso é obtido, simplesmente, invertendo a função matemática do modelo directo definido na equação 8.32, como definido na equação 8.33:

$$\widehat{f}_g(k - d_1(k)) = \frac{-a_1 + \sqrt{a_1^2 - 4a_2 \left( a_0 - \frac{\Delta t(k+1) - p_1 \Delta t(k)}{q_1} \right)}}{2a_2} \quad (8.33)$$

A inversão matematicamente do modelo directo (equação 8.32) só é possível se o parâmetro  $a_2$  for diferente de zero e o valor que se encontra dentro da raiz quadrada for maior que zero.

### Modelo Hammerstein Adaptativo Neuro-Difuso

**Modelo Directo** A função não linear estática é modelizada por um sistema neuro-difuso do tipo Takagi-Sugeno de primeira ordem, como apresentado na equação 8.34:

$$z_g(k - d_1(k)) = \sum_{i=1}^3 \left( \frac{e^{\left( -\frac{(f_g(k - d_1(k)) - c_i)^2}{2\sigma_i^2} \right)}}{\sum_{i=1}^3 \left( e^{\left( -\frac{(f_g(k - d_1(k)) - c_i)^2}{2\sigma_i^2} \right)} \right)} (n_i f_g(k - d_1(k)) + m_i) \right), i = 1..3 \quad (8.34)$$

Este modelo é em tudo idêntico ao apresentado na secção 8.6. Os regressores são os apresentados na equação 8.35:

$$z_g(k - d_1(k)) = f(f_g(k - d_1(k))) \quad (8.35)$$

Neste caso, o modelo neuro-difuso é muito simples, uma vez que só possui uma entrada e uma saída, o que implica que o número de conjuntos difusos seja igual ao número de regras e, neste caso, igual a 3.

O modelo Hammerstein adaptativo neuro-difuso directo do esquentador é definido pela equação 8.36:

$$\widehat{\Delta t}(k+1) = p_1 \Delta t(k) + q_1 z_g(k - d_1(k)) \quad (8.36)$$

em que  $z_g(k - d_1(k))$  está definido na equação 8.34.

O modelo neuro-difuso possui três regras, cujos respectivos consequentes são definidos por dois parâmetros lineares, resultando em seis parâmetros. Quanto aos parâmetros não lineares, o modelo possui uma entrada representada através de três conjuntos difusos definidos por dois parâmetros não lineares, resultando em seis parâmetros não lineares.

Os parâmetros  $p_1$  e  $q_1$  são parâmetros variáveis a identificar em linha, recursivamente, e  $c_i$ ,  $\sigma_i$  e  $n_i$ ,  $m_i$ , com  $i = 1..3$  são os doze parâmetros do modelo neuro-difuso directo a identificar fora de linha.

**Modelo Inverso** Na impossibilidade de inverter a função matemática do modelo da função estática do modelo neuro-difuso directo, o modelo inverso é construído com uma estrutura idêntica à utilizada no modelo directo (número de regras e parâmetros), com a excepção da troca da entrada pela saída e da saída pela entrada do modelo da função estática neuro-difusa directa, como definido na equação 8.37:

$$\widehat{f}_g(k - d_1(k)) = \sum_{i=1}^3 \left( \frac{e^{-\frac{(z_g(k-d_1(k))-c_{il})^2}{2\sigma_{il}^2}}}{\sum_{i=1}^3 e^{-\frac{(z_g(k-d_1(k))-c_{il})^2}{2\sigma_{il}^2}}} (n_i z_g(k - d_1(k))) + m_i \right), i = 1..3 \quad (8.37)$$

O modelo Hammerstein adaptativo neuro-difuso inverso é definido pela equação 8.37, em que  $z_g(k - d_1(k))$  está definido na equação 8.38:

$$z_g(k - d_1(k)) = \frac{\Delta t(k+1) - p_1 \Delta t(k)}{q_1} \quad (8.38)$$

Os parâmetros  $p_1$  e  $q_1$  são os mesmos parâmetros identificados em linha para o modelo Hammerstein adaptativo neuro-difuso directo e  $c_i$ ,  $\sigma_i$  e  $n_i$ ,  $m_i$ , com  $i = 1..3$  são os 12 parâmetros do modelo neuro-difuso da função estática inversa a identificar fora de linha. Apesar de referidos com os mesmos nomes, estes parâmetros tomam valores diferentes dos identificados no modelo neuro-difuso da função estática directa.

### 8.7.2 Identificação do Modelo Hammerstein Adaptativo

A identificação dos modelos Hammerstein propostos para o esquentador é efectuada em duas fases. Numa primeira fase, é identificada, fora de linha, a função não linear estática

utilizando dados estáticos de entrada e saída para um dado caudal fixo. Numa segunda fase, é identificado o modelo de 1ª ordem dinâmico de forma recursiva utilizando dados dinâmicos do sistema.

Como se pretende dividir o modelo do esquentador em dois blocos, o ganho do sistema tem de ser distribuído de forma bem definida pelos dois blocos para evitar redundância nos modelos construídos. Um procedimento muito comum é a imposição de ganho unitário ao modelo linear dinâmico, sendo o restante ganho modelizado pela função não linear estática.

### Identificação do Modelo Hammerstein Adaptativo Polinomial

**Identificação do Modelo Directo** Numa primeira fase é efectuada a identificação da função não linear estática, e numa segunda fase é efectuada a identificação da função linear dinâmica de 1ª ordem.

Considerando o ganho de  $G(\cdot)$  igual a um, os aumentos de temperatura  $\Delta t(k)$  vão corresponder à variável intermédia  $z_g(k)$  em estado estacionário, em que  $z_g(k - d_1(k)) = z_g(k) = \Delta t(k)$ .

Para identificar esta função é necessário um conjunto de dados de entrada e saída em estado estacionário. Este conjunto de dados é retirado do conjunto de dados de treino  $\mathbf{E}^{NE}$  como apresentado na figura 8.17.

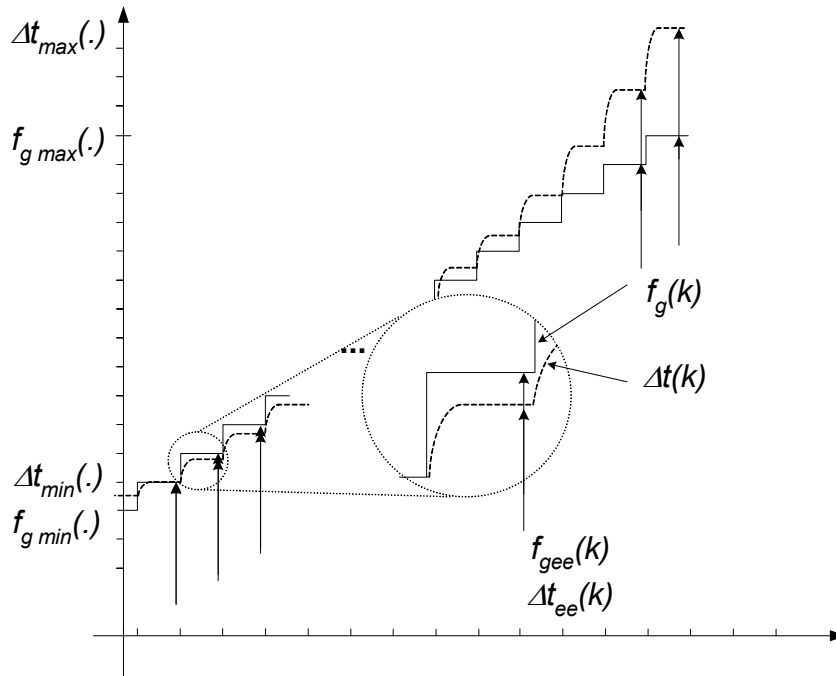


Figura 8.17: Figura ilustrativa do processo de recolha dos dados em estado estacionário  $\mathbf{EE}^{NEE}$  com base nos dados estáticos e dinâmicos  $\mathbf{E}^{NE}$ , para um caudal fixo ( $\Delta t(k)$  - linha tracejada e  $f_g(k)$  - linha contínua).

$$\mathbf{EE}^{NEE} = \{f_{g\ ee}(k); \Delta t_{ee}(k)\}, k = 1..NEE \text{ é o conjunto de dados em estado}$$



estacionário, em que  $f_{gee}(k)$  e de  $\Delta t_{ee}(k)$  são os valores de  $f_g(k)$  e de  $\Delta t(k)$  em estado estacionário e  $NEE$  é o número de elementos de cada vector desse conjunto. O procedimento ilustrado na figura 8.17 retira um par de pontos de  $(f_g(k); \Delta t(k))$  em  $\mathbf{E}^{NE}$  em cada degrau de  $f_g(k)$  em estado estacionário. Deste modo  $NEE$  é igual a o número de degraus de  $f_g(k)$  existentes em  $\mathbf{E}^{NE}$ , desde o valor mínimo até ao valor máximo de  $f_g(k)$ . A identificação dos três parâmetros  $a_0$ ,  $a_1$  e  $a_2$  do polinómio de segunda ordem é realizada, fora de linha, através do método dos mínimos quadrados.

Para identificar a função linear dinâmica de 1ª ordem utilizaram-se os dados de treino  $\mathbf{D}^{ND}$  e o método recursivo de mínimos quadrados (RLS). Deste modo, os parâmetros  $p_1$  e  $q_1$  são identificados, em que a entrada é o sinal intermédio  $z_g(k - d_1(k))$  e a saída é  $\Delta t(k)$ . Os valores de  $z_g(k - d_1(k))$  são calculados com o auxilio da função polinomial identificada na primeira fase.

Estes parâmetros tem de ser identificados em linha uma vez que variam de caudal para caudal. Como o caudal de água possui variações em forma de degrau, estas vão provocar variações em degrau nos parâmetros do modelo linear. Para uma adaptação eficaz da variação rápida dos parâmetros lineares utiliza-se a técnica da reinicialização da matriz das covariâncias (*Reset Covariance Matrix - RCM*). Esta reinicialização é efectuada com base no valor do erro de estimação.

O diagrama de blocos da figura 8.18 mostra o esquema de identificação recursivo utilizado para estimar os parâmetros do modelo linear dinâmico do modelo Hammerstein adaptativo.

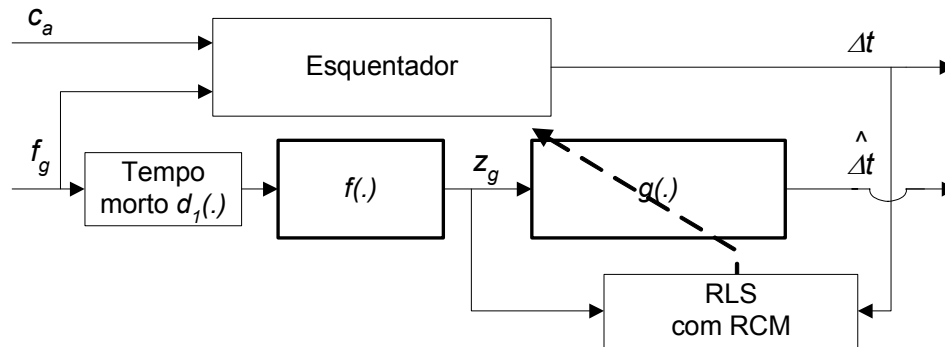


Figura 8.18: Identificação recursiva do bloco linear do modelo Hammerstein.

**Identificação do Modelo Inverso** Dentro das condições já referidas anteriormente, o modelo Hammerstein adaptativo polinomial inverso não necessita de ser identificado, uma vez que é possível inverter, matematicamente, o modelo Hammerstein adaptativo polinomial directo.

### Identificação do Modelo Hammerstein Adaptativo Neuro-Difuso

**Identificação do Modelo Directo** Numa primeira fase, é efectuada a identificação da função não linear estática e, numa segunda fase, é efectuada a identificação da função linear dinâmica de 1ª ordem.

Utilizando os dados em estado estacionário  $\mathbf{EE}^{NEE} = \{f_{gee}(k); \Delta t_{ee}(k)\}$ ,  $k = 1..NEE$ , e a função de custo definida na equação 8.19, o modelo neuro-difuso directo da função não linear estática é identificado. Os doze parâmetros do modelo neuro-difuso  $c_i$ ,  $\sigma_i$  e  $n_i$ ,  $m_i$  com  $i = 1..3$ , são identificados fora de linha, utilizando o método híbrido de identificação referido na identificação do modelo neuro-difuso, apresentado na secção 8.6.

Para identificar a função linear dinâmica de 1ª ordem utilizou-se exactamente o mesmo procedimento utilizado para a identificação do modelo linear de 1ª ordem do modelo Hammerstein adaptativo polinomial.

**Identificação do Modelo Inverso** A identificação do modelo inverso da função não linear estática é realizada da mesma forma que a identificação do modelo directo, com a excepção dos regressores utilizados  $\mathbf{EE}^{NEE} = \{\Delta t_{ee}(k); f_{gee}(k)\}$ ,  $k = 1..NEE$ . A estrutura e o número de parâmetros do modelo neuro-difuso inverso da função estática é igual à do modelo neuro-difuso directo. Os parâmetros do modelo neuro-difuso inverso da função não linear estática, apesar de referidos com nomes idênticos, tomam valores diferentes dos identificados para o modelo neuro-difuso directo.

Os parâmetros do modelo linear inverso de 1ª ordem são os identificados para o modelo linear directo 1ª ordem do modelo Hammerstein.

### 8.7.3 Resultados

#### Modelo Não Linear Estático Polinomial

A figura 8.19 apresenta os resultados obtidos pelo modelo polinomial da função não linear estática.

#### Modelo Linear Dinâmico

A figura 8.20 apresenta a evolução da identificação dos parâmetros  $p_1$  e  $q_1$  do modelo linear de 1ª ordem do modelo Hammerstein adaptativo polinomial. Estes parâmetros variam essencialmente com o caudal de água como era previsto pelos conhecimentos *a priori* existentes sobre o esquentador.

#### Modelo Hammerstein Adaptativo Polinomial

A tabela 8.3 apresenta os valores de  $\sqrt{J(\boldsymbol{\theta}, \mathbf{D}^{ND})}$  e  $V(\boldsymbol{\theta}, \mathbf{H}^{NH})$  obtidos, respectivamente, para os modelos Hammerstein adaptativo polinomial directo e inverso.

Função de Custo	Modelo	
	Directo [°C]	Inverso [%]
Treino - $\sqrt{J(\boldsymbol{\theta}, \mathbf{D}^{ND})}$	0,41	1,21
Teste - $V(\boldsymbol{\theta}, \mathbf{H}^{NH})$	0,40	1,22

Tabela 8.3: MSEs obtidos com o modelo Hammerstein adaptativo polinomial.

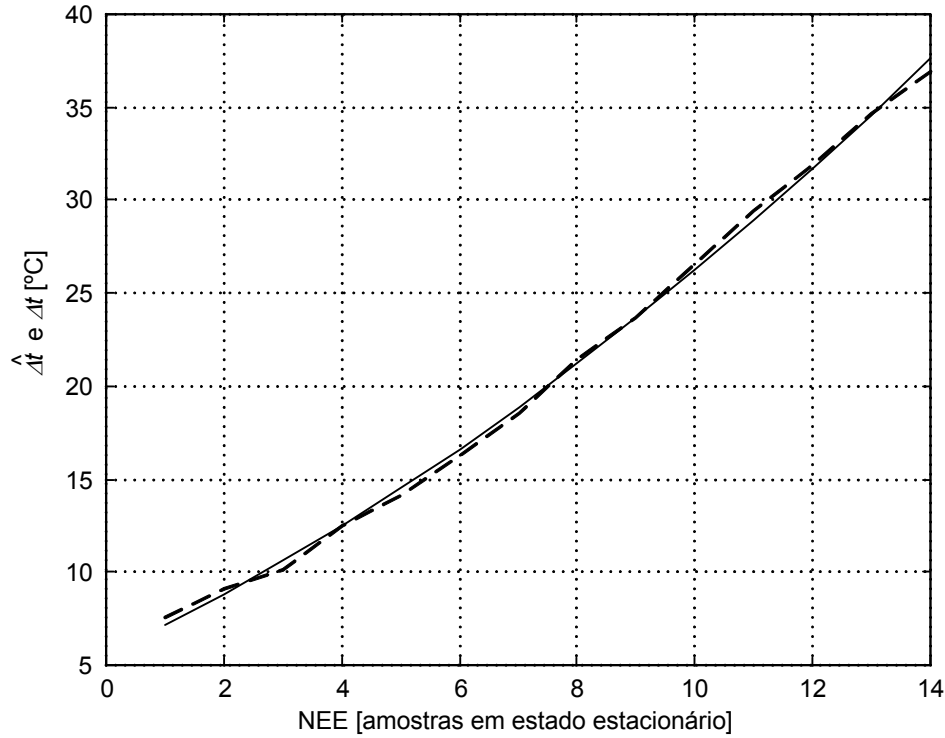


Figura 8.19: Resultados da identificação da função não linear estática do modelo Hammerstein adaptativo polinomial directo, função estática real e estimada (linha tracejada e contínua, respectivamente).

**Modelo Directo** Na figura 8.21 estão apresentados os resultados obtidos pelo modelo directo Hammerstein adaptativo polinomial. De cima para baixo, o gráfico 1 apresenta o aumento da temperatura estimado e real (linha tracejada e contínua, respectivamente) e o gráfico 2 apresenta o erro entre os dois sinais referidos para os dados de treino.

**Modelo Inverso** Os resultados obtidos pelo modelo inverso Hammerstein adaptativo polinomial estão apresentados na figura 8.22. De cima para baixo, o gráfico 1 apresenta o fluxo de gás estimado e real (linha tracejada e contínua, respectivamente) e o gráfico 2 apresenta o erro entre os dois sinais referidos para os dados de treino.

### Modelo Hammerstein Adaptativo Neuro-Difuso

A tabela 8.4 apresenta os valores de  $\sqrt{J(\boldsymbol{\theta}, \mathbf{D}^{ND})}$  e  $V(\boldsymbol{\theta}, \mathbf{H}^{NH})$  obtidos, respectivamente, para os modelos Hammerstein adaptativo neuro-difuso directo e inverso.

Uma vez que os resultados obtidos com o modelo directo e inverso Hammerstein neuro-difuso são muito semelhantes aos apresentados nas figuras 8.21 e 8.22 obtidos utilizando o modelo polinomial, estes não serão apresentados neste documento.

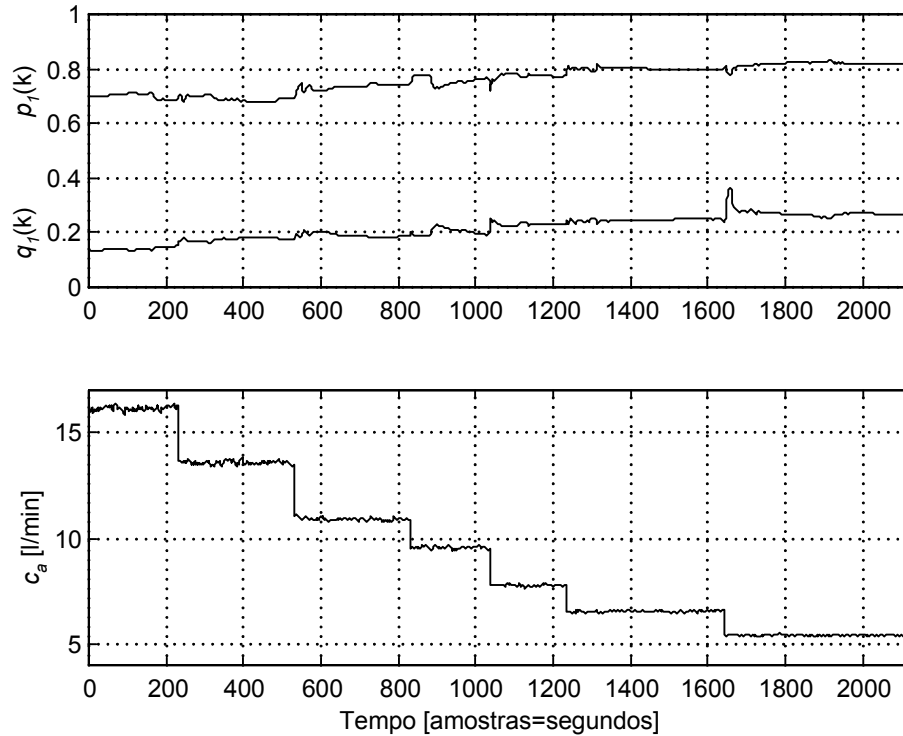


Figura 8.20: Evolução da identificação dos parâmetros do modelo linear para o modelo Hammerstein adaptativo polinomial para os dados de teste. De cima para baixo, gráfico 1 - parâmetros estimados e gráfico 2 - caudal da água.

Função de Custo	Modelo	
	Directo [ $^{\circ}C$ ]	Inverso [%]
Treino - $\sqrt{J(\theta, \mathbf{D}^{ND})}$	0,40	1,19
Teste - $V(\theta, \mathbf{H}^{NH})$	0,41	1,20

Tabela 8.4: MSEs obtidos com o modelo Hammerstein Adaptativo neuro-difuso.

### 8.7.4 Comentários

O modelo Hammerstein do esquentador foi identificado em duas fases. Na primeira fase, são identificados fora de linha os parâmetros da função não linear estática; na segunda fase, são identificados “em linha” os parâmetros do modelo linear dinâmico. O algoritmo de adaptação em linha do modelo Hammerstein aumenta a complexidade matemática do modelo, no entanto, este trás vantagens óbvias de adaptabilidade do modelo a perturbações desconhecidas.

Pelas figuras 8.21 e 8.22 pode observar-se que os sinais reais e estimados pelos modelos adaptativos directo e inverso estão praticamente sobrepostos, o que indicia uma boa modelização.

Os resultados apresentados nas tabelas 8.3 e 8.4 mostram que os modelos Hammerstein adaptativos obtêm erros de modelização baixos.

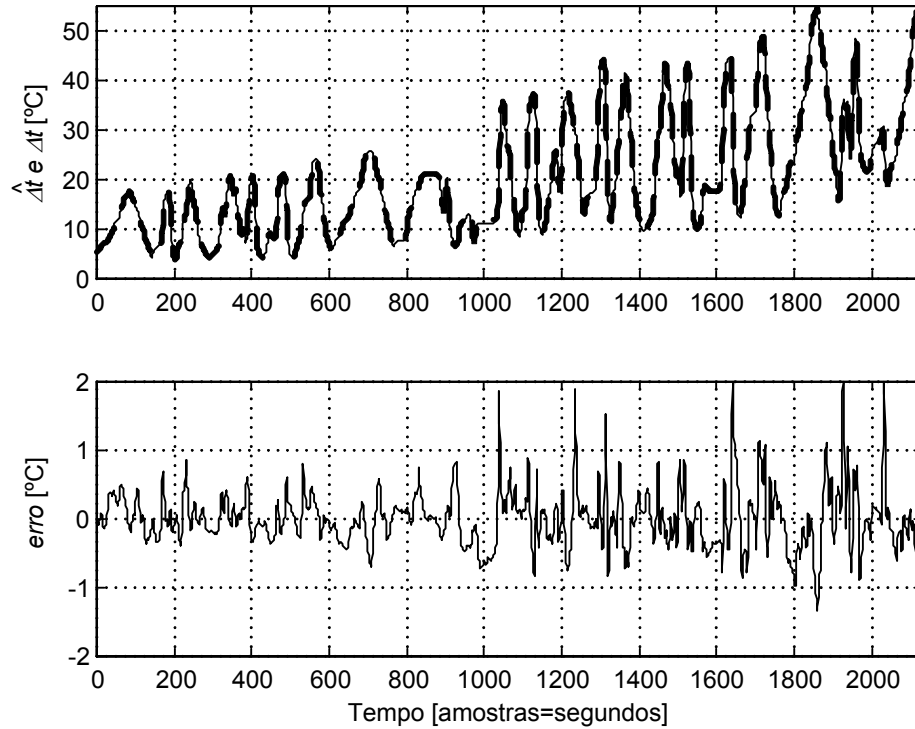


Figura 8.21: Resultados do modelo Hammerstein adaptativo polinomial directo. De cima para baixo, gráfico 1 - aumento da temperatura estimada e real (linha tracejada e contínua, respectivamente) e gráfico 2 - sinal de erro entre os dois sinais anteriores, em ordem ao tempo.

Quanto às duas variantes de modelo Hammerstein apresentadas para o esquentador, o modelo Hammerstein adaptativo polinomial é o “preferível”, uma vez que atinge igualmente bons resultados com uma estrutura muito mais simples. Neste caso, a não linearidade estática do esquentador é correctamente modelizada por um polinómio de ordem dois. Os sistemas neuro-difusos devem ser utilizados quando a função não linear estática é mais complexa como, por exemplo, uma função quadrática com saturação ou uma função não linear com histerese.

## 8.8 Modelização e Identificação Híbrida

Dos diversos métodos de modelização apresentados nesta dissertação, a modelização híbrida é o método que utiliza a maior quantidade de conhecimentos fenomenológicos na construção dos respectivos modelos. A modelização híbrida baseia-se nas equações fenomenológicas permitindo a construção de modelos com elevado grau de interpretabilidade, contudo, o conhecimento que advém das equações fenomenológicas do sistema nem sempre é completo existindo partes menos conhecidas ou mesmo desconhecidas que são, tipicamente, modelizadas por funções do tipo caixa cinzenta ou negra.

A primeira tarefa a realizar na modelização híbrida do esquentador é a definição

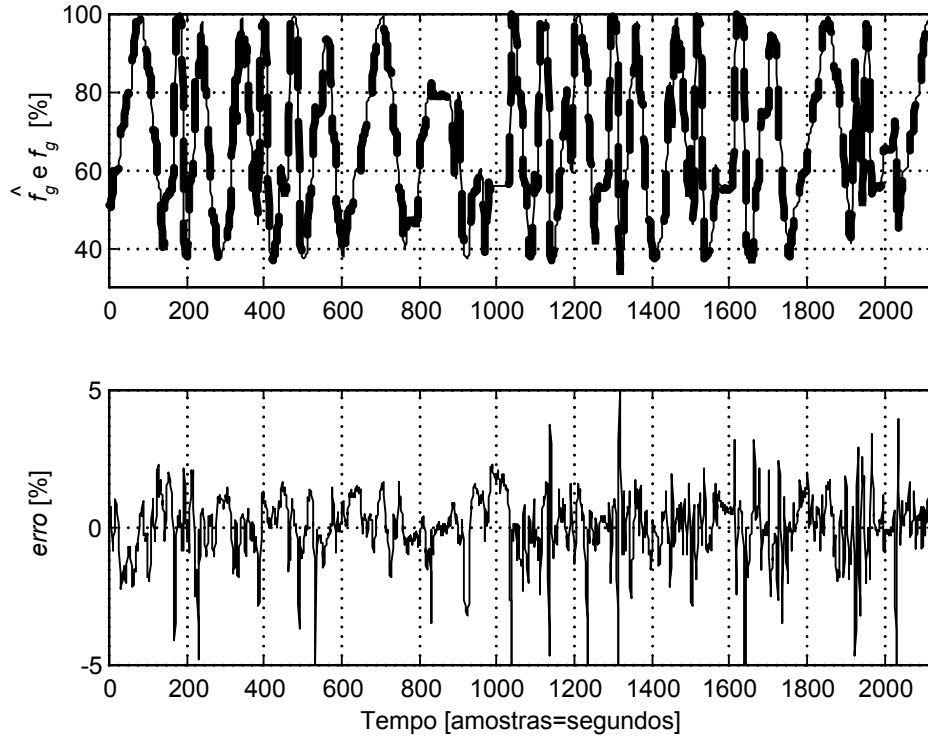


Figura 8.22: Resultados do modelo Hammerstein adaptativo polinomial inverso. De cima para baixo, gráfico 1 - fluxo de gás estimado e real (linha tracejada e contínua, respectivamente) e gráfico 2 - sinal de erro entre os dois sinais anteriores, em ordem ao tempo.

das equações de balanço de energia que irão permitir a definição do respectivo modelo físico.

### 8.8.1 Modelo Físico

O comportamento do esquentador pode ser descrito baseado no balanço macroscópico de energia ou potência térmica ocorrido no sistema. Este balanço de potência térmica é formulado baseado nos princípios de conservação de energia e descrito pelas respectivas equações diferenciais.

A equação de balanço de energia do esquentador está definida na equação 8.39:

$$\frac{de_s(t)}{dt} = p_{ffg}(t) - p_p(t) + p_{af}(t) - p_{aq}(t) \quad (8.39)$$

em que  $de_s(t)/dt$  é a variação da energia do sistema em ordem ao tempo,  $p_{ffg}(t)$  é a potência térmica fornecida pelo fluxo de gás,  $p_p(t)$  é a potência térmica perdida pelo sistema,  $p_{af}(t)$  é a potência térmica que entra no sistema através da água fria e  $p_{aq}(t)$  é a potência térmica que sai do sistema através da água quente.

O esquentador possui, como apresentado na secção 8.3, dois tempos mortos variáveis entre as duas variáveis de entrada e a variável de saída. Sendo  $\tau_{m1}(t)$  o tempo

morto entre as potências térmicas fornecidas e perdas resultantes da aplicação de um determinado fluxo de gás e a variação da energia no sistema e sendo  $\tau_{m2}(t)$  o tempo morto entre as potências térmicas que entram e que saem resultantes da circulação de um determinado caudal de água e a variação da energia no sistema. Assim, o balanço de potência térmica efectivo passa a ser definido pela equação 8.40:

$$\frac{de_s(t)}{dt} = p_{ffg}(t - \tau_{m1}(t)) - p_p(t - \tau_{m1}(t)) + p_{af}(t - \tau_{m2}(t)) - p_{aq}(t - \tau_{m2}(t)) \quad (8.40)$$

Substituindo os valores de  $p_{af}(t - \tau_{m1}(t))$  e  $p_{aq}(t - \tau_{m1}(t))$  pelas respectivas equações físicas, obtém-se a expressão apresentada na equação 8.41:

$$\begin{aligned} \frac{de_s(t)}{dt} = & p_{ffg}(t - \tau_{m1}(t)) - p_p(t - \tau_{m1}(t)) \\ & - c_a(t - \tau_{m2}(t))C_E t_{aq}(t) + c_a(t - \tau_{m2}(t))C_E t_{af}(t) \end{aligned} \quad (8.41)$$

em que  $C_E$  é o calor específico da água.

A substituição de  $p_{ffg}(t - \tau_{m1}(t))$  e  $p_p(t - \tau_{m1}(t))$  pelas respectivas equações físicas não é efectuada uma vez que estas não são completamente conhecidas. No entanto, denominando a diferença entre estas duas potências por potência térmica a absorver pela água  $p_{aa}(t - \tau_{m1}(t))$ , verifica-se que esta é proporcional ao aumento da temperatura da água ( $\Delta t(t) = t_{aq}(t) - t_{af}(t)$ ) e por sua vez, é função estática do fluxo de gás  $f_g(t - \tau_{m1}(t))$ . Substituindo  $p_{ffg}(t - \tau_{m1}(t)) - p_p(t - \tau_{m1}(t))$  por  $p_{aa}(t - \tau_{m1}(t))$  na equação 8.41 obtém-se a equação 8.42:

$$\frac{de_s(t)}{dt} = p_{aa}(t - \tau_{m1}(t)) - c_a(t - \tau_{m2}(t))C_E \Delta t(t) \quad (8.42)$$

Partindo do pressuposto apresentado no capítulo 7, que considera que em cada utilização do esquentador a temperatura da água fria permanece praticamente constante e substituindo a variação da energia no sistema em ordem ao tempo pela respectiva equação física, assumindo que  $dt_{aq}(t) \Leftrightarrow d\Delta t(t)$ , obtém-se a equação de balanço de potência térmica do esquentador:

$$MC_E \frac{d\Delta t(t)}{dt} = p_{aa}(t - \tau_{m1}(t)) - c_a(t - \tau_{m2}(t))C_E \Delta t(t) \quad (8.43)$$

em que  $M$  é a massa de água que se encontra dentro do permutador.

Escrevendo a equação 8.43 no domínio Laplaciano com  $\tau_{m1}(t) = \tau_{m1}$  e  $\tau_{m2}(t) = \tau_{m2}$  valores constantes obtém-se a equação 8.44:

$$P_{aa}(s)e^{-\tau_{m1}s} = MC_E \Delta T(s)s + C_a(s)e^{-\tau_{m2}s}C_E \Delta T(s) \quad (8.44)$$

$$\frac{\Delta T(s)}{P_{aa}(s)} = \frac{1}{MC_E s + C_a(s)e^{-\tau_{m2}s}C_E} e^{-\tau_{m1}s} \quad (8.45)$$

em que  $P_{aa}(s)e^{-\tau_{m1}s}$ ,  $\Delta T(s)$  e  $C_a(s)e^{-\tau_{m2}s}$  são as transformadas de Laplace de  $p_{aa}(t - \tau_{m1})$ ,  $\Delta t(t)$  e  $c_a(t - \tau_{m2})$ , respectivamente.

Após alguma manipulação algébrica, a função de transferência do esquentador (relação entre a variação do aumento da temperatura à saída e potência térmica a absorver pela água) é apresentada na equação 8.46:

$$\frac{\Delta T(s)}{P_{aa}(s)} = \frac{\frac{1}{C_a(s)e^{-\tau_{m2}s}C_E}}{\frac{M}{C_a(s)e^{-\tau_{m2}s}s} + 1} e^{-\tau_{m1}s} \quad (8.46)$$

Rescrevendo a equação 8.46, como apresentado na equação 8.47, é possível efectuar a transformação desta função para o domínio discreto:

$$\frac{\Delta T(s)}{P_{aa}(s)} = \frac{\frac{1}{C_a(s)e^{-\tau_{m2}s}C_E} \frac{C_a(s)e^{-\tau_{m2}s}}{M}}{s + \frac{C_a(s)e^{-\tau_{m2}s}}{M}} e^{-\tau_{m1}s} \quad (8.47)$$

Aplicando o operador  $e^{-sh} = z^{-1}$  à equação 8.47 e substituindo  $\tau_{m1}$  e  $\tau_{m2}$  pelas respectivas aproximações discretas  $d_1$  e  $d_2$ , obtém-se a equação 8.48:

$$\frac{\Delta T(z)}{P_{aa}(z)} = \frac{1}{C_a(z)z^{-d_2}C_E} \frac{1 - e^{-\frac{C_a(z)z^{-d_2}}{M}h}}{z - e^{-\frac{C_a(z)z^{-d_2}}{M}h}} z^{-d_1} \quad (8.48)$$

Baseado na equação 8.48 e com  $h = 1$  s, e sabendo que  $d_1$  e  $d_2$  variam, ou seja  $d_1(k)$  e  $d_2(k)$ , a equação 8.49 representa a função transferência do esquentador escrita do domínio discreto, cuja saída estimada é designada por  $\widehat{\Delta t}(k+1)$ :

$$\begin{aligned} \widehat{\Delta t}(k+1) &= \left( e^{-\frac{c_a(k-d_2(k))}{M}} \right) \Delta t(k) + \\ &\left( \frac{1}{c_a(k-d_2(k))C_E} \left( 1 - e^{-\frac{c_a(k-d_2(k))}{M}} \right) \right) p_{aa}(k-d_1(k)) \end{aligned} \quad (8.49)$$

O modelo físico do esquentador escrito no domínio discreto (equação 8.49) possui uma parte desconhecida em que os conhecimentos físicos são escassos. Essa parte menos conhecida é a potência térmica a absorver pela água  $p_{aa}(k-d_1(k))$ . Segundo o estudo realizado sobre o comportamento estático do esquentador, apresentado no capítulo 7, verifica-se que a potência térmica a absorver pela água (proporcional ao aumento da temperatura da água) depende do fluxo de gás aplicado ao queimador de uma forma estática e não linear (forma quadrática). Para modelizar esta parte menos conhecida do modelo físico do esquentador, é utilizada uma função do tipo caixa cinzenta ou negra  $f_{CCN}(\cdot)$ , como se apresenta na equação 8.50:

$$p_{aa}(k-d_1(k)) = f_{CCN}(f_g(k-d_1(k))) \quad (8.50)$$

Deste modo, o modelo final do esquentador é constituído por duas funções: a função que representa o modelo físico e a função  $f_{CCN}(\cdot)$  que representa a parte desconhecida do modelo global, como apresentado na equação 8.51:

$$\begin{aligned} \widehat{\Delta t}(k+1) &= \left( e^{-\frac{c_a(k-d_2(k))}{M}} \right) \Delta t(k) + \\ &\left( \frac{1}{c_a(k-d_2(k))C_E} \left( 1 - e^{-\frac{c_a(k-d_2(k))}{M}} \right) \right) (f_{CCN}(f_g(k-d_1(k)))) \end{aligned} \quad (8.51)$$



### 8.8.2 Modelo Híbrido Série

A modelização baseada na combinação em série de um modelo desconhecido com um modelo fenomenológico é conhecida por modelização híbrida série. Com base no modelo físico definido na equação 8.51 é possível verificar que este possui uma função desconhecida que será modelizada por um modelo do tipo caixa cinzenta ou negra  $f_{CCN}(\cdot)$ . A saída deste modelo  $f_{CCN}(\cdot)$  é uma das entradas do modelo físico como ilustrado na figura 8.23. Assim, o modelo do esquentador é constituído pelo modelo da função desconhecida e pelo modelo físico ligados em série.

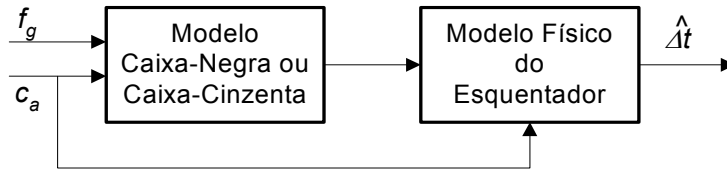


Figura 8.23: Modelo híbrido série do esquentador.

A parte desconhecida ou menos conhecida do sistema resulta do conhecimento incompleto de um dos termos da equação de balanço de energias, equação 8.40.

Esta secção apresenta duas formas diferentes para a modelização da função desconhecida do modelo híbrido série do esquentador. Em primeiro lugar, a referida função é modelizada utilizando um polinómio dando origem ao modelo híbrido série polinomial; em segundo, utilizando um sistema neuro-difuso dando origem ao modelo híbrido série neuro-difuso.

A modelização da função desconhecida utilizando sistemas neuro-difusos é mais geral, uma vez que estes permitem a modelização de qualquer função não linear sem qualquer tipo de restrições apesar de imporem um maior número de parâmetros. A modelização da referida função utilizando funções polinomiais pode ser um pouco limitativa.

#### Modelo Híbrido Série Polinomial

**Modelo Directo** O modelo híbrido série polinomial é constituído pelo modelo físico, apresentado na secção anterior, e por uma função polinomial  $f_{CCN}(\cdot)$ , que modeliza a parte menos conhecida do esquentador.

Baseado nos conhecimentos existentes sobre a componente menos conhecida do modelo global, verifica-se que é suficiente uma função polinomial de ordem 2, definida na equação 8.52, para a modelização de  $f_{CCN}(\cdot)$ :

$$p_{aa}(k - d_1(k)) = f_{CCN}(f_g(k - d_1(k))) \quad (8.52)$$

$$p_{aa}(k - d_1(k)) = a_0 + a_1 f_g(k - d_1(k)) + a_2 f_g(k - d_1(k))^2$$

Deste modo, o modelo híbrido série polinomial é definido pela equação 8.53:

$$\begin{aligned}\widehat{\Delta t}(k+1) &= \left(e^{-\frac{c_a(k-d_2(k))}{M}}\right) \Delta t(k) + \\ &\left(\frac{1}{c_a(k-d_2(k))C_E} \left(1 - e^{-\frac{c_a(k-d_2(k))}{M}}\right)\right) (a_0 + a_1 f_g(k-d_1(k)) + a_2 f_g(k-d_1(k))^2)\end{aligned}\quad (8.53)$$

com  $M = 0,42$  Kg (massa de água dentro do permutador) e  $C_E = 4186$  JKg<sup>-1</sup>K<sup>-1</sup> (calor específico da água).

Definindo as funções  $K_1(k-d_2(k))$  e  $K_2(k-d_2(k))$  apresentadas na equação 8.54:

$$\begin{aligned}K_1(k-d_2(k)) &= \left(e^{-\frac{c_a(k-d_2(k))}{M}}\right) \\ K_2(k-d_2(k)) &= \left(\frac{1}{c_a(k-d_2(k))C_E} \left(1 - e^{-\frac{c_a(k-d_2(k))}{M}}\right)\right)\end{aligned}\quad (8.54)$$

O modelo híbrido série polinomial do esquentador é definido de uma forma mais simples e clara pela equação 8.55:

$$\widehat{\Delta t}(k+1) = K_1(k-d_2(k))\Delta t(k) + K_2(k-d_2(k)) (a_0 + a_1 f_g(k-d_1(k)) + a_2 f_g(k-d_1(k))^2) \quad (8.55)$$

em que os parâmetros  $M$  e  $C_E$  são grandezas físicas conhecidas e  $a_0$ ,  $a_1$  e  $a_2$  são parâmetros desconhecidos que é necessário identificar.

**Modelo Inverso** O modelo inverso é obtido invertendo a função matemática do modelo directo (equação 8.55) como definido na equação 8.56:

$$\widehat{f}_g(k-d_1(k)) = \frac{-a_1 + \sqrt{a_1^2 - 4a_2 \left(a_0 - \frac{\Delta t(k+1) - K_1(k-d_2(k))\Delta t(k)}{K_2(k-d_2(k))}\right)}}{2a_2} \quad (8.56)$$

Só é possível inverter matematicamente o modelo directo se o valor que se encontra dentro da raiz quadrada for maior que zero e o parâmetro  $a_2$  for diferente de zero.

### Modelo Híbrido Série Neuro-Difuso

**Modelo Directo** No modelo híbrido série neuro-difuso, a função desconhecida é modelizada por um sistema do tipo Takagi-Sugeno de primeira ordem, em tudo idêntico ao apresentado na secção 8.6, como é apresentado na equação 8.57:

$$p_{aa}(k-d_1(k)) = \sum_{i=1}^3 \left( \frac{e^{\left(-\frac{(f_g(k-d_1(k))-c_i)^2}{2\sigma_i^2}\right)}}{\sum_{i=1}^3 \left(e^{\left(-\frac{(f_g(k-d_1(k))-c_i)^2}{2\sigma_i^2}\right)}\right)} (n_i f_g(k-d_1(k)) + m_i) \right), i = 1..3 \quad (8.57)$$

Sendo o modelo híbrido série neuro-difuso directo do esquentador é definido pela equação 8.58:

$$\widehat{\Delta t}(k+1) = K_1(k-d_2(k))\Delta t(k) + K_2(k-d_2(k))p_{aa}(k-d_1(k)) \quad (8.58)$$

em que  $p_{aa}(k-d_1(k))$  está definida na equação 8.57.

O modelo neuro-difuso possui três regras cujos respectivos consequentes são definidos por dois parâmetros lineares, resultando num total de seis parâmetros lineares ( $n_i$ ,  $m_i$ , com  $i = 1..3$ ). Quanto aos parâmetros não lineares, o modelo possui uma entrada representada através de três conjuntos difusos diferentes definidos por dois parâmetros não lineares, resultando num total de seis parâmetros não lineares ( $c_i$  e  $\sigma_i$  com  $i = 1..3$ ).

Os parâmetros  $M$  e  $C_E$  são parâmetros físicos conhecidos e  $c_i$ ,  $\sigma_i$  e  $n_i$ ,  $m_i$ , com  $i = 1..3$  são os doze parâmetros do modelo neuro-difuso directo que é necessário identificar.

**Modelo Inverso** Na impossibilidade de inverter a função matemática do modelo desconhecido neuro-difuso directo (equação 8.57), o modelo inverso é construído com uma estrutura idêntica à utilizada no modelo directo (igual número de regras e parâmetros), com a excepção da troca da entrada pela saída e saída pela entrada do modelo neuro-difuso directo, como apresentado na equação 8.59:

$$\widehat{f}_g(k-d_1(k)) = \sum_{i=1}^3 \left( \frac{e^{\left(-\frac{(p_{aa}(k-d_1(k))-c_{il})^2}{2\sigma_{il}^2}\right)}}}{\sum_{i=1}^3 e^{\left(-\frac{(p_{aa}(k-d_1(k))-c_{il})^2}{2\sigma_{il}^2}\right)}}} (n_i p_{aa}(k-d_1(k))) + m_i \right), i = 1..3 \quad (8.59)$$

O modelo híbrido série neuro-difuso inverso é definido pela equação 8.59, em que  $p_{aa}(k-d_1(k))$  esta definida na equação 8.60:

$$p_{aa}(k-d_1(k)) = \frac{\Delta t(k+1) - K_1(k-d_2(k))\Delta t(k)}{K_2(k-d_2(k))} \quad (8.60)$$

Os parâmetros  $M$  e  $C_E$  são parâmetros físicos conhecidos e  $c_i$ ,  $\sigma_i$  e  $n_i$ ,  $m_i$ , com  $i = 1..3$ , são os 12 parâmetros do modelo neuro-difuso inverso a identificar que apesar de referidos com nomes idênticos, estes parâmetros tomam estes valores diferentes dos identificados no modelo neuro-difuso directo.

### 8.8.3 Identificação do Modelo Híbrido Série

#### Modelo Híbrido Série Polinomial

**Identificação do Modelo Directo** O diagrama de blocos do esquema de identificação do modelo polinomial da função desconhecida está ilustrado na figura 8.24, sendo semelhante ao proposto no capítulo 5.

A identificação dos parâmetros do modelo polinomial é efectuada recorrendo ao método dos mínimos quadrados, utilizando os dados de treino  $\mathbf{D}^{ND}$  e minimizando a

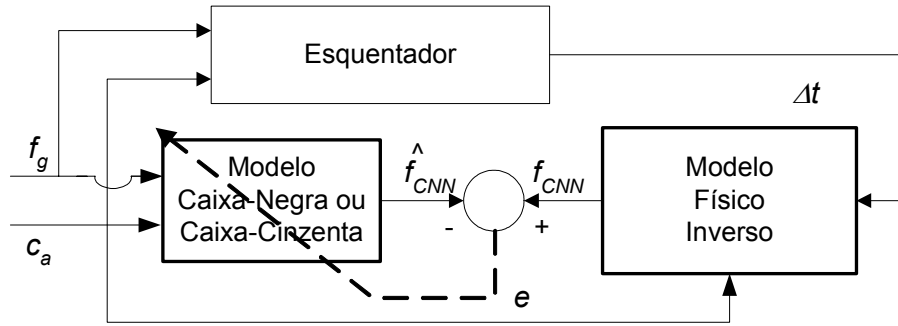


Figura 8.24: Identificação do modelo híbrido série directo.

função de custo  $J(\theta, \mathbf{D}^{ND})$  (equação 8.19) em que o erro  $e(k)$  é dado pela equação 8.61:

$$e(k) = f_{CNN}(f_g(k - d_1(k))) - \hat{f}_{CNN}(f_g(k - d_1(k))) \quad (8.61)$$

$$e(k) = \left( \frac{\Delta t(k+1) - K_1(k - d_2(k)) \Delta t(k)}{K_2(k - d_2(k))} \right) - \hat{f}_{CNN}(f_g(k - d_1(k)))$$

Os parâmetros  $M$  e  $C_E$  são parâmetros físicos e  $a_0$ ,  $a_1$  e  $a_2$  são os parâmetros a identificar.

**Identificação do Modelo Inverso** Como apresentado atrás, o modelo híbrido série polinomial inverso não necessita de ser identificado, uma vez que é possível inverter o modelo directo, como referido para o modelo Hammerstein adaptativo polinomial inverso.

### Modelo Híbrido Série Neuro-Difuso

**Identificação do Modelo Directo** A identificação dos parâmetros do modelo neuro-difuso directo da função desconhecida é efectuada de modo idêntico ao apresentado na secção 8.6, para o modelo neuro-difuso directo global, em que o erro  $e(k)$  é dado pela equação 8.61, estando o diagrama de blocos do esquema de identificação ilustrado na figura 8.24.

**Identificação do Modelo Inverso** A identificação dos parâmetros do modelo neuro-difuso inverso da função desconhecida é efectuada através de um modelo neuro-difuso idêntico ao apresentado para o modelo neuro-difuso directo e baseado no erro  $e(k)$  definido na equação 8.62, em que  $\hat{f}_{CNN}^{-1}(f_g(k - d_1(k)))$  é dado pela equação 8.59:

$$e(k) = f_g(k - d_1(k)) - \hat{f}_{CNN}^{-1}(f_g(k - d_1(k))) \quad (8.62)$$

O respectivo diagrama de blocos do esquema de identificação está ilustrado na figura 8.25.

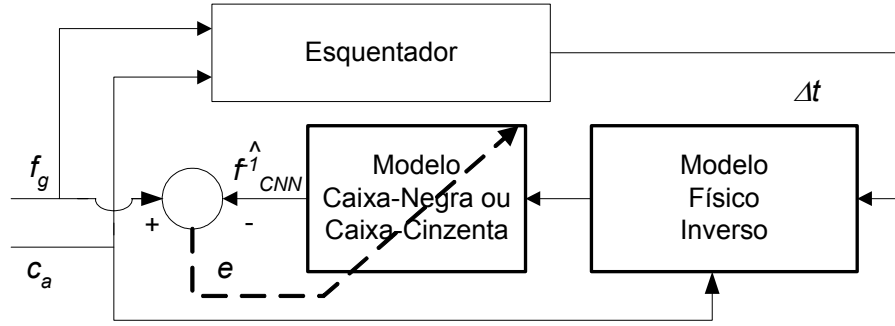


Figura 8.25: Identificação do modelo híbrido série inverso.

Os parâmetros do modelo neuro-difuso inverso, apesar de referidos com nomes idênticos, tomam valores diferentes dos identificados para o modelo neuro-difuso directo da função desconhecida.

### 8.8.4 Resultados

#### Modelo Físico Variável

A figura 8.26 apresenta o cálculo das funções  $K_1(k - d_2(k))$  e  $K_2(k - d_2(k))$  que representam os parâmetros do modelo linear físico do modelo híbrido série. Estes parâmetros variam com o caudal de água como definido na equação 8.54.

#### Modelo Híbrido Série Polinomial

A tabela 8.5 apresenta os valores de  $\sqrt{J(\theta, \mathbf{D}^{ND})}$  e  $V(\theta, \mathbf{H}^{NH})$  obtidos, respectivamente, para os modelos híbrido série polinomial directo e inverso.

Função de Custo	Modelo	
	Directo [ $^{\circ}C$ ]	Inverso [%]
Treino - $\sqrt{J(\theta, \mathbf{D}^{ND})}$	0,60	1,86
Teste - $V(\theta, \mathbf{H}^{NH})$	0,68	1,87

Tabela 8.5: MSEs obtidos com o modelo híbrido série polinomial.

**Modelo Directo** Na figura 8.27 estão apresentados os resultados obtidos pelo modelo directo híbrido série polinomial. De cima para baixo, o gráfico 1 apresenta o aumento da temperatura estimado e real (linha tracejada e contínua, respectivamente) e o gráfico 2 apresenta o erro entre os dois sinais referidos, utilizando os dados de treino.

**Modelo Inverso** Os resultados obtidos pelo modelo inverso híbrido série polinomial estão apresentados na figura 8.28. De cima para baixo, o gráfico 1 apresenta o fluxo de gás estimado e real (linha tracejada e contínua, respectivamente) e o gráfico 2 apresenta o erro entre os dois sinais referidos para os dados de treino.

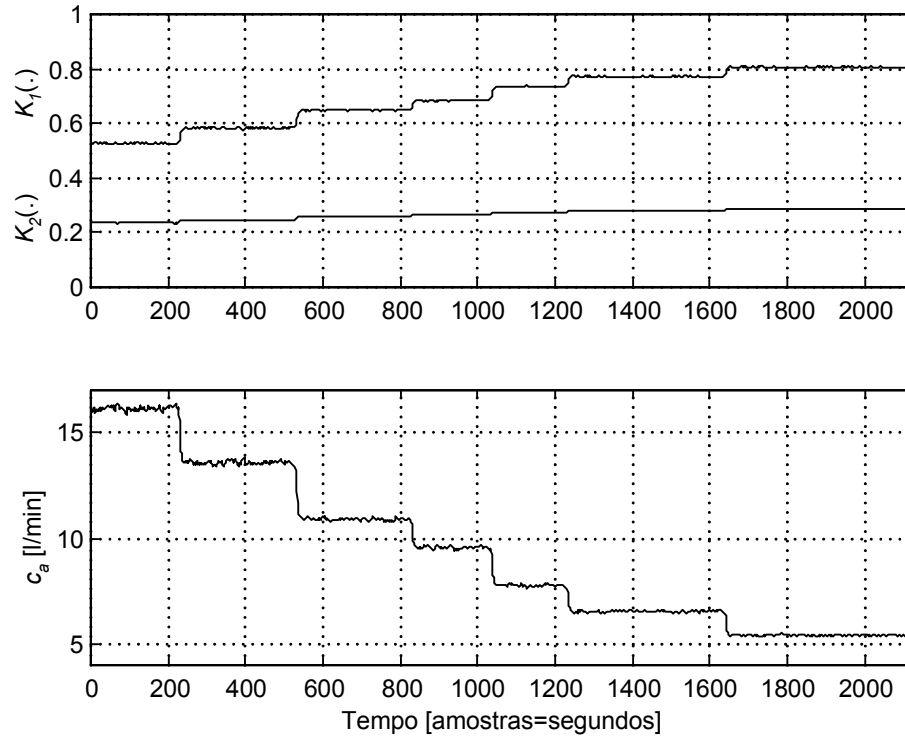


Figura 8.26: Parâmetros do modelo físico linear para o modelo híbrido para os dados de treino. De cima para baixo, gráfico 1 - parâmetros do modelo físico e gráfico 2 - caudal da água.

### Modelo Híbrido Série Neuro-Difuso

A tabela 8.6 apresenta os valores de  $\sqrt{J(\theta, \mathbf{D}^{ND})}$  e  $V(\theta, \mathbf{H}^{NH})$  obtidos, respectivamente, com os modelos híbrido série neuro-difuso directo e inverso.

Função de Custo	Modelo	
	Directo [ $^{\circ}C$ ]	Inverso [%]
Treino - $\sqrt{J(\theta, \mathbf{D}^{ND})}$	0,60	1,86
Teste - $V(\theta, \mathbf{H}^{NH})$	0,67	1,87

Tabela 8.6: MSEs obtidos com o modelo híbrido série neuro-difuso.

Uma vez que os resultados obtidos com o modelo directo e inverso híbrido neuro-difuso são muito semelhantes aos apresentados nas figuras 8.27 e 8.28 obtidos utilizando o modelo polinomial, estes não serão apresentados neste documento.

### 8.8.5 Melhoramentos ao Modelo Híbrido Série

Depois de uma análise detalhada dos erros de modelização obtidos com os modelos híbridos série, verifica-se que estes não estão centrados na recta erro igual a zero (ver

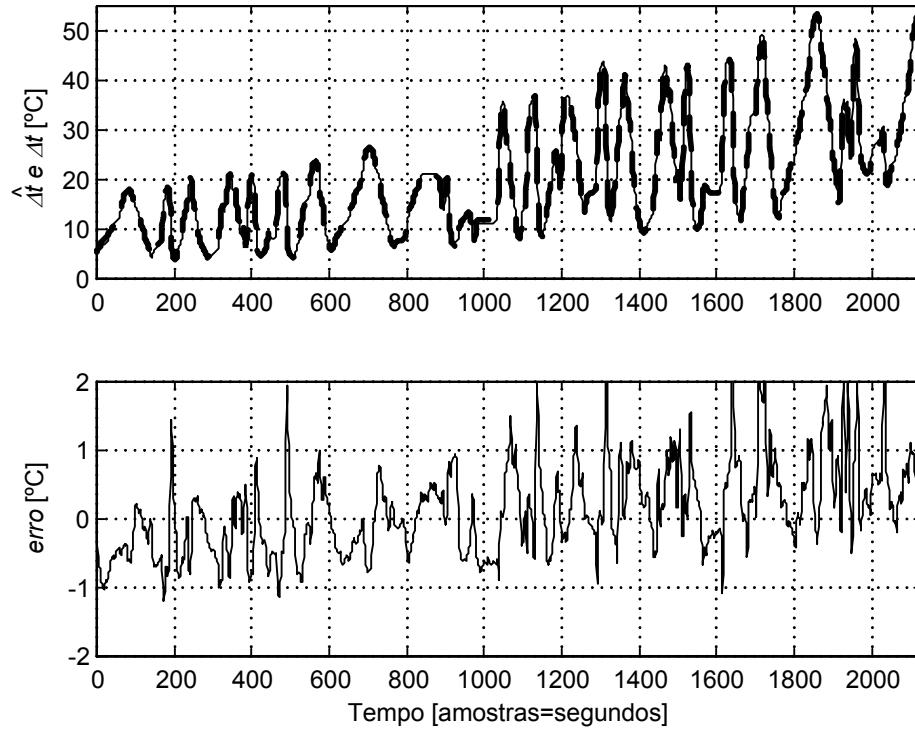


Figura 8.27: Resultados do modelo híbrido série polinomial directo. De cima para baixo, gráfico 1 - aumento da temperatura estimada e real (linha tracejada e contínua, respectivamente) e gráfico 2 - sinal de erro entre os dois sinais anteriores, em ordem ao tempo.

figuras 8.27 e 8.28). Verifica-se ainda que os erros da variável aumento da temperatura são directamente proporcionais à variável aumento da temperatura e que os erros da variável fluxo de gás são inversamente proporcionais à variável aumento da temperatura. Este facto indicia que a potência térmica a absorver pela água possa depender também do aumento da temperatura da água.

Como apresentado em [VM05] a resolução deste problema passa por incluir na função desconhecida  $f_{CCN}(\cdot)$  mais um regressor no modelo da função desconhecida: o aumento da temperatura da água  $\Delta t(k)$ , sendo o novo modelo definido pela equação 8.63:

$$p_{aa}(k - d_1(k)) = f_{CCN}(f_g(k - d_1(k)), \Delta t(k)) \quad (8.63)$$

Este procedimento permite eliminar o problema referido, todavia a complexidade dos modelos aumenta consideravelmente. De facto, os erros de modelização dos modelos híbridos série com e sem a inclusão da variável  $\Delta t(k)$  na função desconhecida não são muito diferentes, por isso, os modelos apresentados nesta dissertação são os modelos mais simples (sem o regressor  $\Delta t(k)$ ).

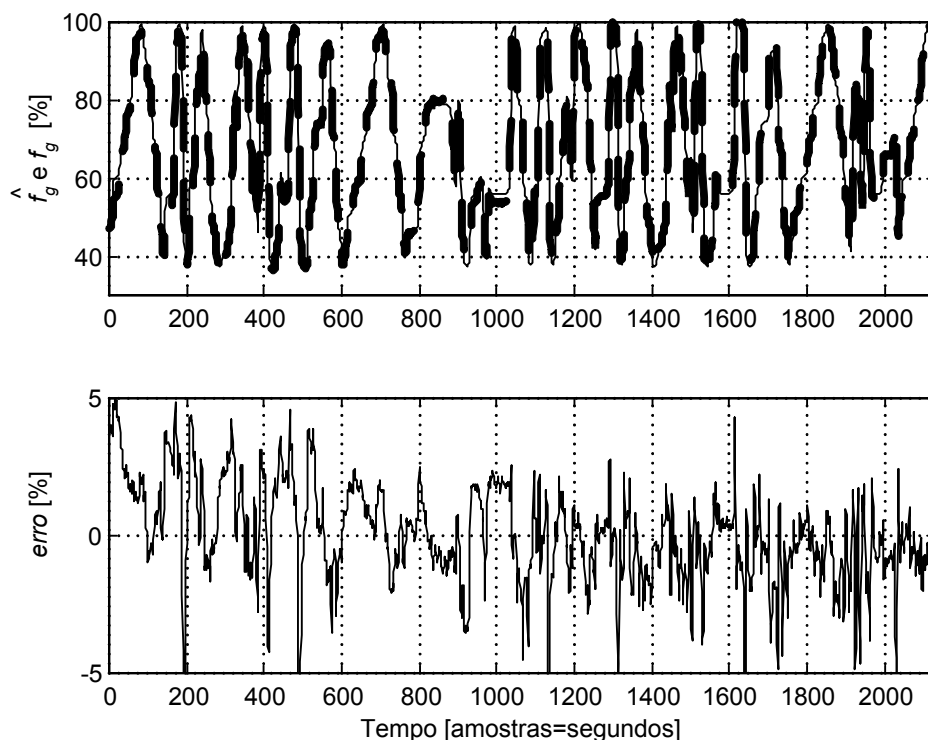


Figura 8.28: Resultados do modelo híbrido série polinomial inverso. De cima para baixo, gráfico 1 - fluxo de gás estimado e real (linha tracejada e contínua, respectivamente) e gráfico 2 - sinal de erro entre os dois sinais anteriores, em ordem ao tempo.

### 8.8.6 Comentários

Pelas figuras 8.27 e 8.28 pode observar-se que os sinais reais e estimados pelos modelos híbridos série directo e inverso estão praticamente sobrepostos o que indicia uma boa modelização.

Os resultados apresentados nas tabelas 8.5 e 8.6 mostram que os modelos híbridos apresentam erros de modelização baixos. A complexidade dos modelos é relativamente baixa e a sua interpretabilidade é elevada.

Quanto às duas variantes de modelos híbridos, pode afirmar-se que o modelo híbrido série polinomial é preferível, uma vez que atinge igualmente bons resultados com uma estrutura mais simples. De notar que a modelização da função desconhecida utilizando um polinómio pode torna-se limitativa, visto que a definição correcta deste modelo vai exigir uma grande quantidade de conhecimentos acerca das variáveis envolvidas, das suas relações e do fenómeno “desconhecido” a modelizar. Quando a função a modelizar for completamente desconhecida, a modelização neuro-difusa será a mais indicada, já que permite a aproximação de qualquer tipo de função sendo a sua identificação baseada simplesmente num conjunto de dados de entrada e saída.



## 8.9 Comparação de resultados

A tabela 8.7 apresenta os valores de  $RMSE$  obtidos com os dados de treino e teste para os quatro modelos identificados: modelo linear (ML), modelo neuro-difuso (MND), modelo Hammerstein adaptativo polinomial (MHAP) e modelo híbrido série polinomial (MHSP).

Modelo	Directo [ $^{\circ}C$ ]	Inverso [%]
ML - Treino	4,50	11,00
ML - Teste	4,90	11,50
MND - Treino	0,36	1,17
MND - Teste	0,59	1,28
MHAP - Treino	0,41	1,21
MHAP - Teste	0,40	1,22
MHSP - Treino	0,60	1,86
MHSP - Teste	0,68	1,87

Tabela 8.7: MSEs obtidos pelos quatro modelos: linear, neuro-difuso, Hammerstein adaptativo polinomial e híbrido série polinomial.

Como se pode observar pela tabela 8.7 os erros obtidos pelos modelos lineares são, como era de esperar, muito maiores do que os erros obtidos pelos restantes modelos.

Os erros alcançados pelos três modelos cinzentos não lineares são semelhantes, sendo os erros obtidos pelos modelos híbridos sensivelmente maiores, uma vez que a sua estrutura matemática é mais limitada (o número de parâmetros do modelo físico é limitado pelas equações físicas que foram definidas).

Através da observação das funções matemáticas dos modelos do esquentador apresentadas neste capítulo pode concluir-se que a sua complexidade está intimamente ligada ao seu nível de interpretabilidade: quanto menor for a complexidade matemática dos modelos maior é o seu nível de interpretabilidade.

O modelo neuro-difuso do esquentador é caracterizado por um nível de interpretabilidade relativamente baixo e uma estrutura matemática relativamente complexa. O modelo Hammerstein possui níveis de interpretabilidade e complexidade matemática intermédios. Finalmente, o modelo híbrido é caracterizado por um nível de interpretabilidade relativamente elevado e uma complexidade matemática relativamente baixa.

## 8.10 Conclusão

Neste capítulo foram apresentados o procedimento de aquisição de dados, o modelo discreto do esquentador (modelo dos tempos mortos e modelo sem tempos mortos), as estruturas de treino e, finalmente, os modelos directos e inversos utilizando quatro metodologias de modelização: linear, neuro-difusa, Hammerstein adaptativa e híbrida série.

Os modelos lineares do esquentador apresentam estruturas matemáticas simples mas erros de modelização muito elevados.

Os modelos neuro-difusos apresentam um erro de modelização baixo, possuem níveis de interpretabilidade reduzidos e estruturas de complexidade matemática relativamente elevada. O tempo dispendido e o nível de conhecimento *a priori* necessários para a construção deste tipo de modelos é relativamente pequeno.

Os modelos Hammerstein adaptativos apresentam erros de modelização baixos e níveis de interpretabilidade bastante razoáveis. Contudo, a identificação em linha do modelo linear dinâmico sobrecarrega computacionalmente a utilização deste modelo.

Os modelos híbridos série apresentam erros de modelização baixos, complexidades matemáticas relativamente baixas e níveis de interpretabilidade bastante elevados. O tempo dispendido e o nível de conhecimento *a priori* necessários para a construção deste tipo de modelos é relativamente elevado.

Os erros de modelização obtidos pelos três modelos não lineares são muito semelhantes, no entanto, a complexidade matemática e a interpretabilidade destes variam significativamente com a abordagem utilizada. Perante os vários métodos de modelização abordados, a modelização híbrida polinomial é a abordagem que permite a construção de modelos com menor complexidade matemática e com maior interpretabilidade.

# Capítulo 9

## Controlo do Esquentador

### 9.1 Introdução

Neste capítulo são apresentados os resultados do controlo do esquentador resultantes da aplicação de quatro malhas de controlo. Estas malhas são baseadas no controlador Smith preditivo, apresentado no capítulo 6, utilizando os quatro modelos referidos no capítulo anterior. São deste modo definidas as condições de teste e avaliado o desempenho de quatro controladores: controlador Smith preditivo linear (CSPL), controlador Smith preditivo neuro-difuso (CSPND), controlador Smith preditivo Hammerstein adaptativo (CSPHA) polinomial e controlador Smith preditivo híbrido série (CSPHS) polinomial. O controlador Smith preditivo linear é apresentado apenas para comparação de resultados com os obtidos com os controladores não lineares.

Ao nível do controlo baseado nos modelos neuro-difusos implementaram-se duas abordagens diferentes: uma que considera o tempo morto fixo [VM02] e outra o tempo morto variável [VM03b]. Nesta dissertação é apresentado o CSPND tomando em consideração os dois tempos mortos variáveis do sistema.

Relativamente ao controlo baseado em modelos orientados por blocos construíram-se dois controladores Smith preditivos Hammerstein adaptativos [VM04a], em que as funções estáticas não lineares dos respectivos modelos foram modelizadas à custa de uma função polinomial e de um sistema neuro-difuso. Nesta dissertação é apresentado o CSPHA que permite simplificar a malha de controlo Smith preditiva utilizando modelos lineares adaptativos (linearização do sistema relativamente ao fluxo de gás) tendo em conta o tempo morto relativo ao fluxo de gás.

Ao nível do controlo baseado em modelos híbridos testaram-se três controladores. No primeiro controlador, o modelo híbrido foi definido com base em conhecimentos físicos e identificado utilizando algoritmos genéticos [VM04e], no segundo controlador, a parte desconhecida do modelo físico foi definida utilizando uma função polinomial de ordem dois [VM04f] e no terceiro, a parte desconhecida foi modelizada através de um modelo neuro-difuso [VDM05]. Nesta dissertação é apresentado o CSPHS em que o bloco desconhecido do modelo híbrido foi modelizado com uma função polinomial. Este controlador tem em conta os dois tempos mortos do sistema e permite também simplificar a malha de controlo do esquentador (linearização do sistema relativamente ao fluxo de gás).

Finalmente, são apresentadas as conclusões analisando a raiz dos erros quadráticos médios (*RMSEs*), as sobrelevações/subelevações relativas às variações do sinal de referência, os tempos de estabelecimento relativos às variações do caudal de água e do sinal de referência e finalmente os tempos de execução dos vários algoritmos.

Neste capítulo encontram-se mais duas das contribuições apresentadas neste trabalho. A definição da malha de controlo Smith preditiva para modelos com múltiplas entradas com tempos mortos diferentes e a definição das malhas de controlo Hammerstein adaptativa polinomial e híbrida série polinomial, em que é efectuada a linearização do sistema relativamente a uma das variáveis de entrada (fluxo de gás) simplificando o modelo e o algoritmo de controlo.

## 9.2 Condições de Teste

Para testar e avaliar as várias malhas de controlo propostas neste capítulo, foram definidos os perfis para as variáveis temperatura desejada (sinal de referência  $r$ ) e caudal de água  $c_a$ .

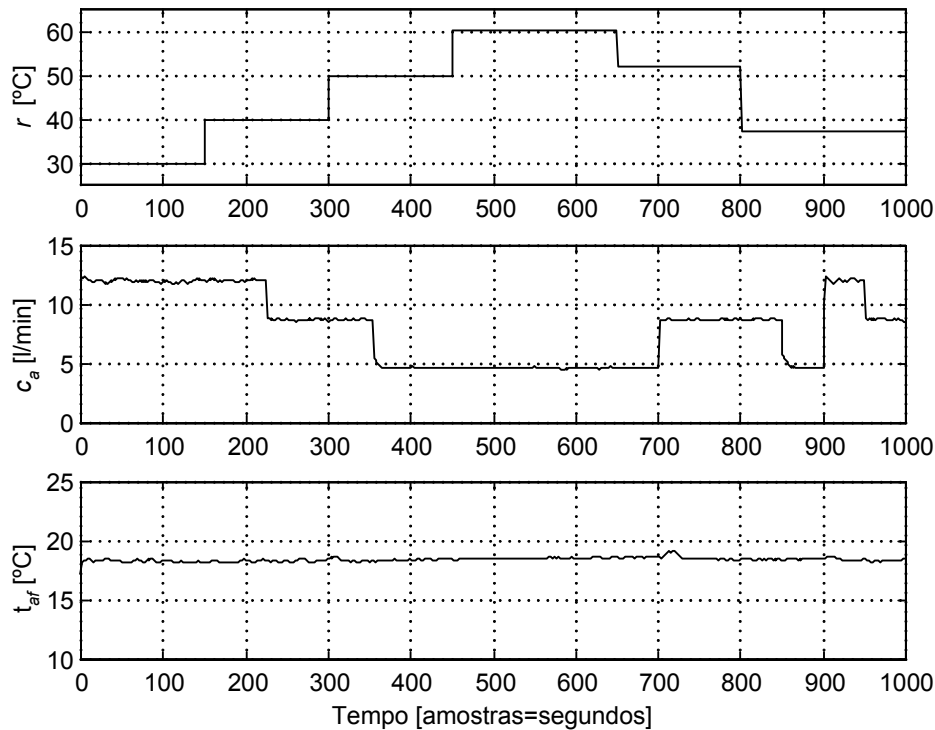


Figura 9.1: Perfis de teste, de cima para baixo, gráfico 1 -  $r(k)$  sinal de referência para a temperatura da água quente, gráfico 2 -  $c_a(k)$  caudal de água e gráfico 3 -  $t_{af}(k)$  temperatura da água fria.

A variável temperatura da água fria  $t_{af}$ , é uma variável não controlada que permaneceu praticamente constante durante a realização dos testes (18,2°C).

Os perfis definidos e ilustrados na figura 9.1 para o teste e avaliação dos controladores, garantem que o esquentador irá funcionar dentro da sua gama de potências estáticas admissíveis, ou seja, numa potência estática contida no intervalo  $[80; 325]$  [Kcal/min]. No caso em que seja exigido o funcionamento do esquentador a uma potência estática fora deste intervalo, o controlador irá saturar no máximo ou no mínimo do sinal de controlo, apresentando erro em estado estacionário.

A definição dos perfis de teste das variáveis temperatura de referência e caudal de água é baseada nos valores que as variáveis tomam habitualmente durante o funcionamento do esquentador. As variações definidas nestes perfis são um pouco mais exigentes (muitas e grandes variações) do que as variações típicas verificadas durante o funcionamento do esquentador. Com estes perfis a maioria das zonas de funcionamento são testadas.

Numa utilização típica do esquentador, o sinal de referência mantém-se constante, existindo a possibilidade de ocorrerem variações em forma de degrau no caudal de água (variações de dois a três l/min que correspondem a perturbações provocadas pela abertura de outras torneiras). No caso da utilização do esquentador para um duche doméstico, a temperatura de referência típica estará fixa num valor entre  $37^{\circ}\text{C}$  e  $42^{\circ}\text{C}$  para o caudal de água definido pelo utilizador. Claro que outros modos de funcionamento serão possíveis daí a exigência imposta nas variações dos perfis de teste apresentados na figura 9.1.

O caudal de água depende do caudal definido pelo utilizador e da pressão da rede de água disponibilizada. O caudal de água ilustrado na figura 9.1 apresenta variações relativamente fortes. No entanto, as variações típicas desta variável são bastante mais “suaves” (uma resposta eficiente do controlador a estas variações garante uma resposta igualmente eficiente para variações menores).

Para avaliar o índice de desempenho dos vários controladores utilizou-se a raiz quadrada do erro quadrático médio (*Root Mean Square Error* - *RMSE*) definido na equação 9.1:

$$RMSE = \sqrt{\frac{1}{N} \sum_{k=1}^N e_t(k)^2} \quad (9.1)$$

em que  $N$  é o número de amostras de  $e_t(k)$  (erro entre o sinal de referência  $r(k)$  e a temperatura real da água quente  $t_{aq}(k)$ ).

Para avaliar a complexidade computacional dos vários controladores, necessariamente dependente da complexidade matemática dos modelos, foram medidos os tempos de execução dos respectivos algoritmos executados no ambiente Matlab num computador pessoal Celeron da Intel a 400 MHz com 192 Mbytes de RAM.

## 9.3 Escolha do Controlador

O esquentador é um sistema não linear com tempo morto variável do qual se possuem modelos directos e inversos definidos e identificados no capítulo 8. Com base nesta

informação, várias malhas de controlo baseadas em modelos podem ser utilizadas para o seu controlo automático.

A escolha da malha de controlo foi efectuada com base nos seguintes critérios: utilização de um controlador baseado em modelo, algoritmos de controlo computacionalmente simples, robustos e que permitam respeitar as especificações de desempenho definidas no capítulo 7.

Das várias malhas de controlo baseadas em modelos, apresentadas no capítulo 6, destacam-se duas pela sua adequação à maioria dos requisitos exigidos: controlo Smith predictivo (capítulo 6, secção 6.4) e controlo predictivo generalizado (capítulo 6, secção 6.5). Estas duas malhas de controlo permitem a construção de controladores robustos e eficazes baseados em modelos não lineares para sistemas com tempo morto variável. No entanto, relativamente à complexidade matemática estas malhas apresentam diferenças muito significativas. Perante as características destas duas malhas de controlo e perante os critérios de escolha definidos, a malha seleccionada para o controlo do esquentador foi a malha Smith predictiva.

## 9.4 Controlador Smith Preditivo

A malha de controlo Smith predictiva utilizada foi baseada na malha de controlo Smith predictiva genérica, apresentada no capítulo 6. Além da utilização do modelo inverso como controlador, esta malha necessita de um modelo global do esquentador sem os tempos mortos e de um modelo global do esquentador com os respectivos tempos mortos.

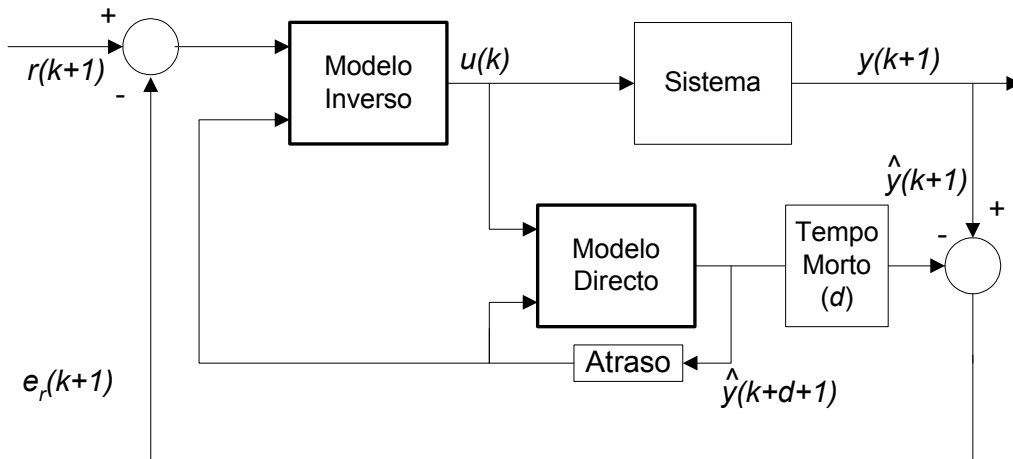


Figura 9.2: Controlador Smith Preditivo com modelo inverso para sistemas SISO de 1ª ordem.

Para sistemas SISO a introdução do tempo morto do sistema no modelo pode ser efectuada através da colocação do seu modelo antes ou depois do modelo do sistema. Pelo diagrama de blocos da malha de controlo Smith predictiva ilustrada na figura 9.2

observa-se que é possível obter os dois sinais de saída do modelo com e sem o tempo morto utilizando um único modelo em que o tempo morto é colocado depois do modelo sem tempo morto. No entanto, para sistemas MISO como é o caso do esquentador, os tempos mortos relativos às várias entradas não podem ser colocados depois do modelo do sistema, uma vez que, os vários tempos mortos podem ser diferentes.

O modelo do esquentador, definido no capítulo 8, é constituído pelos blocos com os tempos mortos relativos às duas entradas do sistema seguidos do bloco com o modelo sem tempos mortos. A introdução em série dos blocos com os tempos mortos depois do modelo do esquentador não corresponde ao modelo correcto do esquentador.

Na definição da malha de controlo Smith preditiva, os dois sinais de saída do modelo com e sem tempos mortos, não podem ser obtidos utilizando um único modelo directo como apresentado no diagrama de blocos da figura 9.2.

A correcta malha de controlo terá de conter conforme ilustrado no diagrama de blocos da figura 9.3 dois modelos directos separados: um sem tempos mortos (Bloco 1) e outro com tempos mortos (Bloco 2).

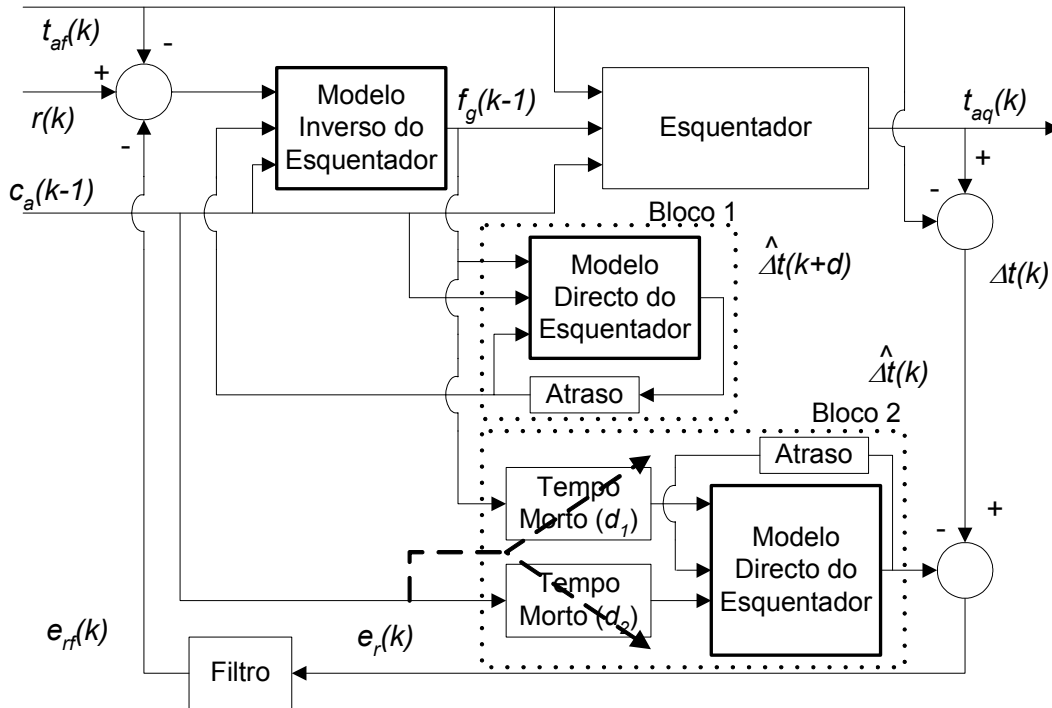


Figura 9.3: Controlador Smith preditivo do esquentador (sistema MISO que utiliza os regressores típicos utilizados num sistema de 1º ordem).

O modelo directo sem tempo morto (Bloco 1) é o responsável pela realimentação do aumento da temperatura avançada para o modelo inverso permitindo a predição do sinal de controlo. O modelo com os dois tempos mortos colocados antes do modelo sem tempos mortos (Bloco 2) é responsável pela realimentação do erro entre o modelo directo e o sistema real.

A execução dos dois modelos directos em separado permite obter, correctamente, as duas saídas do modelo com e sem tempos mortos necessárias para a correcta implementação desta malha de controlo.

A saída estimada do modelo directo sem tempo morto é  $\hat{\Delta}t(k+d)$ , em que  $d = d_1(k)$  ou  $d_2(k)$  referente ao tempo morto relativo ao fluxo de gás ou ao caudal de água.

A variável a controlar é a temperatura da água quente. Todavia, os modelos do esquentador utilizam o aumento da temperatura da água como variável de saída, e não, a temperatura da água quente. Deste modo, a malha de controlo é responsável, como ilustrado na figura 9.3, pelo cálculo do aumento da temperatura desejado a aplicar no modelo inverso  $\Delta t(k) = r(k) - t_{af}(k)$  e pelo aumento da temperatura real  $\Delta t(k) = t_{aq}(k) - t_{af}(k)$  a utilizar na comparação com a saída estimada pelo modelo directo  $\hat{\Delta}t(k)$ .

Para garantir o emparelhamento entre os modelos inverso e directo do esquentador, o regressor utilizado é o aumento da temperatura estimada e não o aumento da temperatura real.

A malha de controlo possui ainda um filtro linear passa-baixo do tipo  $e_{rf}(k) = \alpha e_r(k) + (1 - \alpha)(e_{rf}(k - 1))$  com  $0 < \alpha \leq 1$ , de forma a aumentar a sua robustez. A constante de tempo deste filtro é definida através do parâmetro  $\alpha$ .

O período de amostragem utilizado foi de um segundo sendo, naturalmente igual ao período de amostragem, definido no capítulo 7, para a aquisição dos dados de entrada e saída do sistema.

De seguida são apresentadas as quatro malhas de controlo implementadas e os respectivos resultados. Estas malhas de controlo são avaliadas analisando as raízes quadradas dos erros quadráticos médios (*RMSEs*), as sobrelevações/subelevações relativas às variações do sinal de referência, os tempos de estabelecimento relativos às variações do caudal de água e do sinal de referência e finalmente os tempos de execução dos algoritmos.

## 9.5 Controlador Smith Preditivo Linear

### 9.5.1 Malha de Controlo

O diagrama de blocos do controlador Smith preditivo linear (CSPL) está ilustrado na figura 9.4. Esta malha de controlo, baseada na malha apresentada na figura 9.3, utiliza o modelo linear do esquentador.





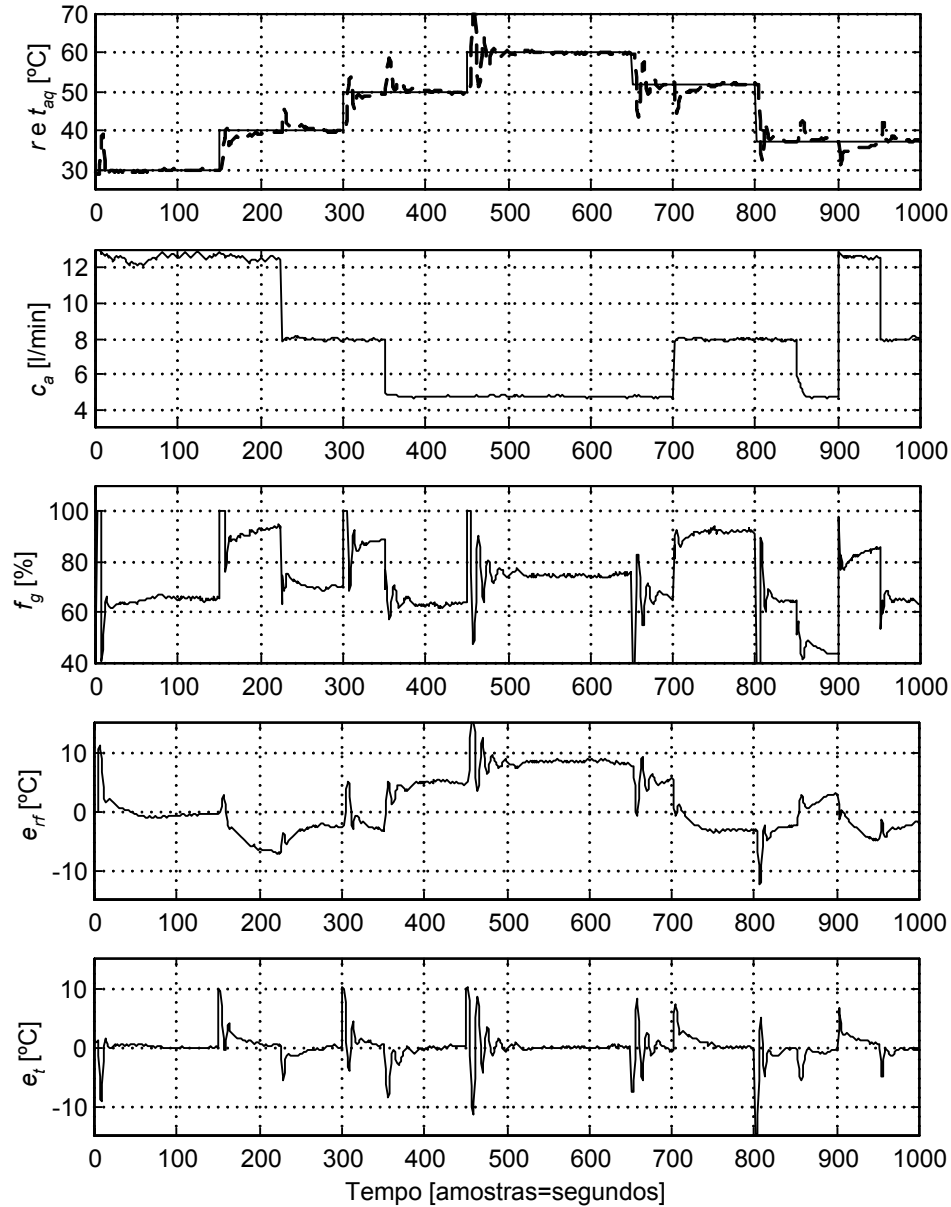


Figura 9.5: Resultados de controlo com o CSPL.

Como se pode observar pela figura 9.5, o sinal de controlo é muito activo, principalmente nas variações da referência, provocando oscilações na temperatura de saída. Este comportamento surge devido ao forte desajuste entre o modelo linear e o próprio esquentador.

É visível ainda a não linearidade do sistema, uma vez que para iguais aumentos da temperatura de referência as respostas apresentam tempos de subida/descida e de estabelecimento diferentes.

Os resultados de controlo obtidos com este controlador apresentam uma raiz quadrada do erro quadrático médio de  $2,28^{\circ}C$ .

O tempo de execução de cada iteração do algoritmo de controlo Smith preditivo

linear é de  $1,50 \times 10^{-3}$  segundos.

### 9.5.3 Comentários

Pela figura 9.5 pode observar-se que o sinal de controlo apresenta alguma oscilação provocando alguma perturbação na temperatura de saída. Apesar de tudo, a malha de controlo Smith preditiva consegue controlar o esquentador anulando o erro em estado estacionário, uma vez que o modelo directo e inverso estão emparelhados.

Os resultados de controlo apresentam, relativamente às variações do sinal de referência, sobreelevações/subelevações que variam de 3 a 10°C, verificando-se um aumento destes valores para caudais baixos. Os valores obtidos ultrapassam sempre valor máximo admitido de 1°C. Os resultados de controlo, relativamente às variações do caudal de água e do sinal de referência, apresentam tempos de estabelecimento que variam de 14 a 40 segundos sendo quase sempre superiores ao valor máximo admitido de 15 segundos.

A observação da evolução do sinal  $e_{rf}(k)$  permite verificar que o modelo e o esquentador respondem de forma relativamente diferente. Esta diferença provoca a degradação dos resultados de controlo.

## 9.6 Controlador Smith Preditivo Neuro-Difuso

### 9.6.1 Malha de Controlo

A figura 9.6 ilustra o diagrama de blocos do controlador Smith preditivo neuro-difuso (CSPND). Neste caso, o modelo do esquentador utilizado é o modelo neuro-difuso apresentado no capítulo 8.

Uma vez que o erro de modelização obtido pelo modelo neuro-difuso é muito baixo é de esperar que o controlador baseado neste modelo apresente um bom desempenho permitindo atingir as especificações de controlo.

A constante de tempo do filtro passa-baixo utilizado nesta malha de controlo é de 0,95 s ( $\alpha = 0,65$ ). Este filtro aumenta a robustez do controlador “suavizando” o sinal de controlo.

### 9.6.2 Resultados

Os resultados obtidos com o controlador Smith preditivo neuro-difuso, aplicando os mesmos perfis de teste, estão ilustrados na figura 9.7. De cima para baixo, são apresentados, no gráfico 1 - a temperatura desejada (linha contínua) e a temperatura real (linha tracejada), no gráfico 2 - o caudal de água, no gráfico 3 - o fluxo de gás aplicado ao queimador, no gráfico 4 - o erro  $e_{rf}(k)$  entre o aumento da temperatura real e estimada e no gráfico 5 - o erro  $e_t(k)$  entre a referência e a temperatura de saída.

Pela figura 9.7 pode observar-se que o sinal de controlo não apresenta grandes oscilações, compensando as variações dos sinais de teste sem provocar sobreelevações/subelevações excessivas na temperatura da água quente.

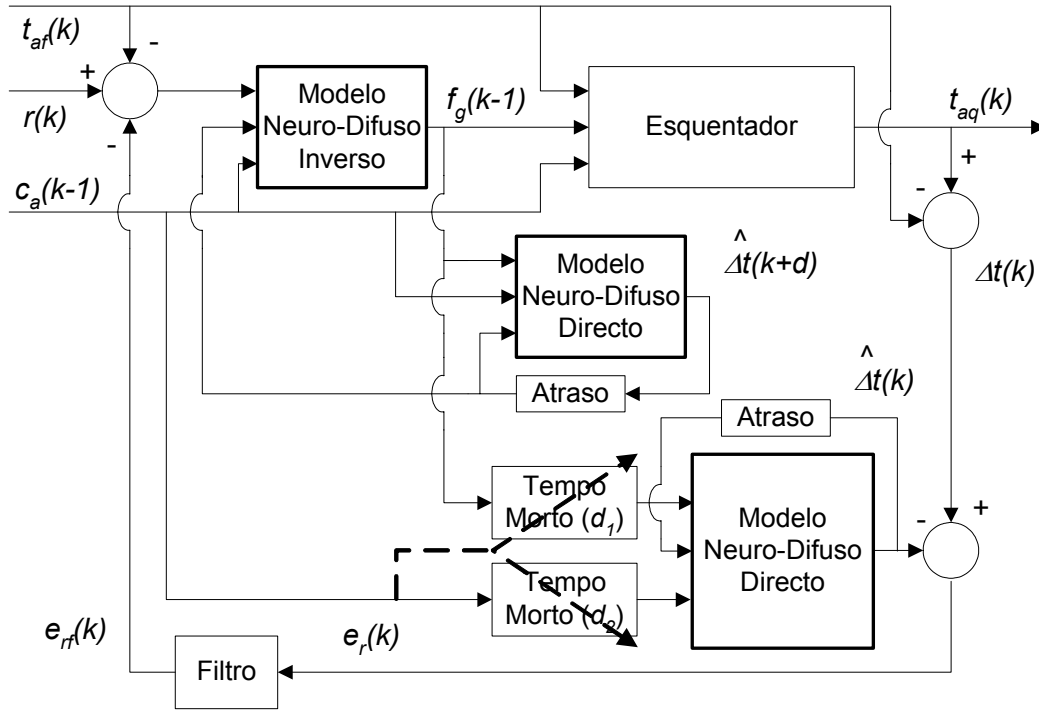


Figura 9.6: Controlador Smith preditivo neuro-difuso.

Os resultados de controlo obtidos com o controlador Smith preditivo neuro-difuso apresentam uma raiz quadrada do erro quadrático médio de  $1,72^{\circ}\text{C}^2$ . O tempo de execução de cada iteração deste algoritmo de controlo é de  $3,62 \times 10^{-3}$  segundos.

### 9.6.3 Comentários

Esta malha de controlo consegue respeitar os critérios de controlo exigidos. Os resultados deste controlador apresentam, relativamente às variações do sinal de referência, sobrelevações/subelevações que variam de  $0,1$  a  $0,9^{\circ}\text{C}$ , obtendo sempre valores inferiores ao valor máximo admitido. Relativamente às variações do caudal de água e da referência, os resultados de controlo apresentam tempos de estabelecimento que variam de 4 a 10 segundos sendo inferiores ao valor máximo admitido de 15 segundos.

Os resultados de controlo respeitam os requisitos de controlo e permitem obter um erro quadrático médio “baixo”.

A observação da evolução do sinal  $e_{rf}(k)$  permite verificar que o modelo e o esquentador respondem de forma muito semelhante, uma vez que a diferença entre a saída do modelo e do esquentador é praticamente nula.

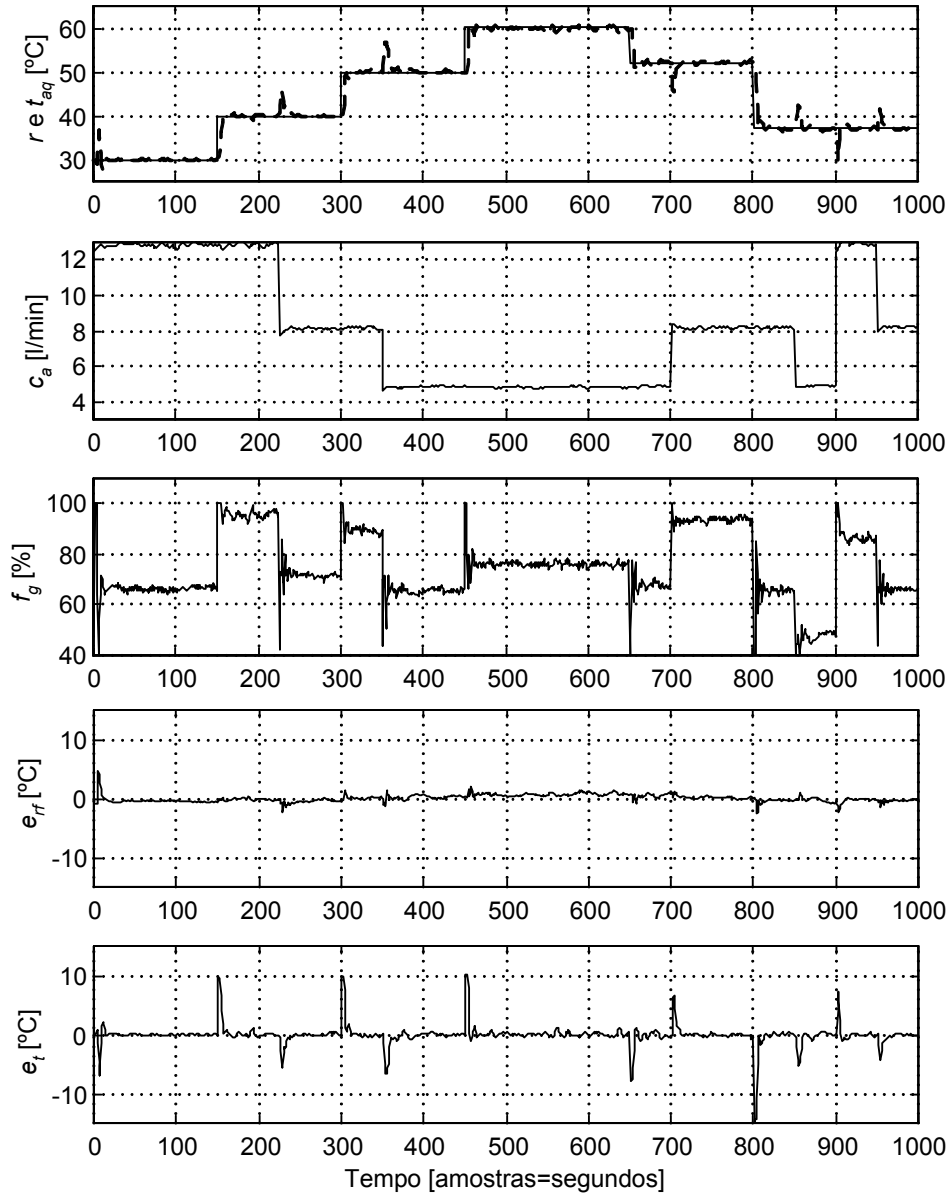


Figura 9.7: Resultados de controlo com o CSPND.

## 9.7 Controlador Smith Preditivo Hammerstein

### 9.7.1 Malha de Controlo

O diagrama de blocos do controlador Smith preditivo Hammerstein adaptativo (CS-PHA) polinomial está ilustrado na figura 9.8.

Conforme referido no capítulo 8, este modelo não utiliza o caudal da água como variável de entrada exigindo adaptação em linha. Como neste caso o modelo do esquentador só tem uma entrada e uma saída, é possível trocar a ordem dos blocos, como referido na secção 9.4, (bloco com o tempo morto e bloco com o modelo sem

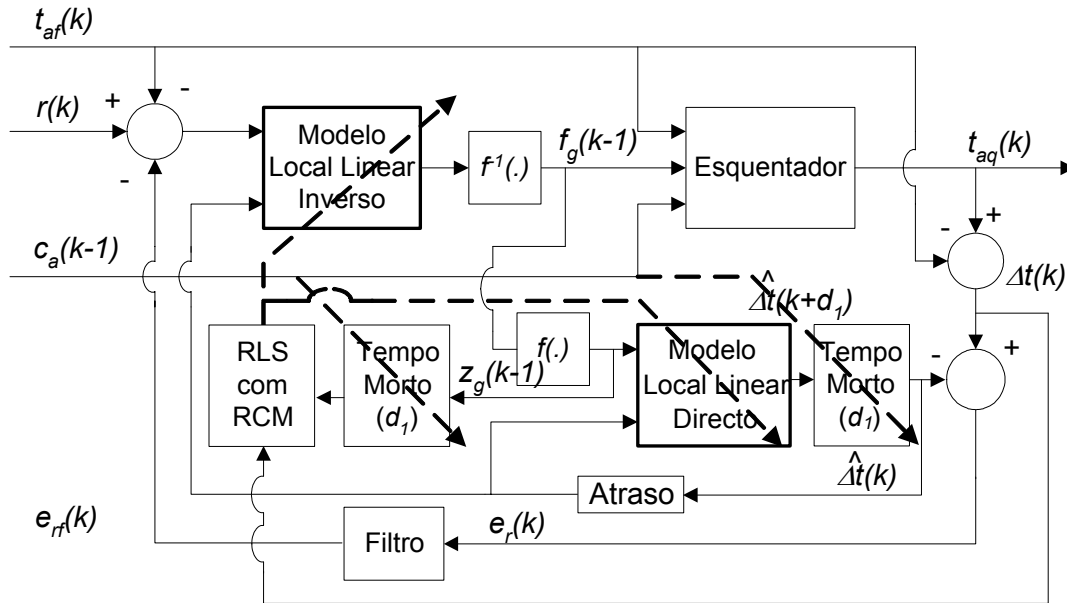


Figura 9.8: Controlador Smith preditivo Hammerstein adaptativo polinomial.

tempo morto) sem cometer erros de modelização. Esta troca apresentada na figura 9.8 permite simplificar a malha de controlo.

O bloco RLS (*Recursive Least Squares*) com RCM (*Reset Covariance Matrix*) é o bloco responsável pela adaptação dos parâmetros dos modelos lineares directo e inverso (o modelo inverso é obtido invertendo matematicamente o modelo directo). Este bloco executa o algoritmo recursivo do método dos mínimos quadrados com reinicialização da matriz das covariâncias para erros de estimação acima de um determinado valor pré-definido, acelerando a adaptação dos parâmetros (fundamentalmente nas variações de caudal de água).

Apesar do modelo Hammerstein adaptativo polinomial corresponder à série de uma função  $f(\cdot)$  não linear estática, com um modelo linear dinâmico variável  $g(\cdot)$  (figura 9.8), este esquema de controlo pode ser simplificado, como se ilustra na figura 9.9, eliminando as funções não lineares estáticas  $f(\cdot)$  e  $f^{-1}(\cdot)$  dos modelos directo e inverso. Estas funções podem ser eliminadas, como se mostra no diagrama de blocos da figura 9.9 (linearização do sistema relativamente ao fluxo de gás), desde que o sinal de controlo seja transformado por  $f^{-1}(\cdot)$  antes de ser aplicado à válvula de gás.

Neste caso o modelo do esquentador é um modelo linear dinâmico de 1ª ordem cujo ganho e constante de tempo variam de forma desconhecida com o caudal de água. Para adaptar este modelo aos diversos caudais possíveis é efectuada uma adaptação em linha dos seus parâmetros.

A constante de tempo do filtro passa-baixo utilizado neste controlador é idêntica à utilizada no filtro do controlador Smith preditivo neuro-difuso.



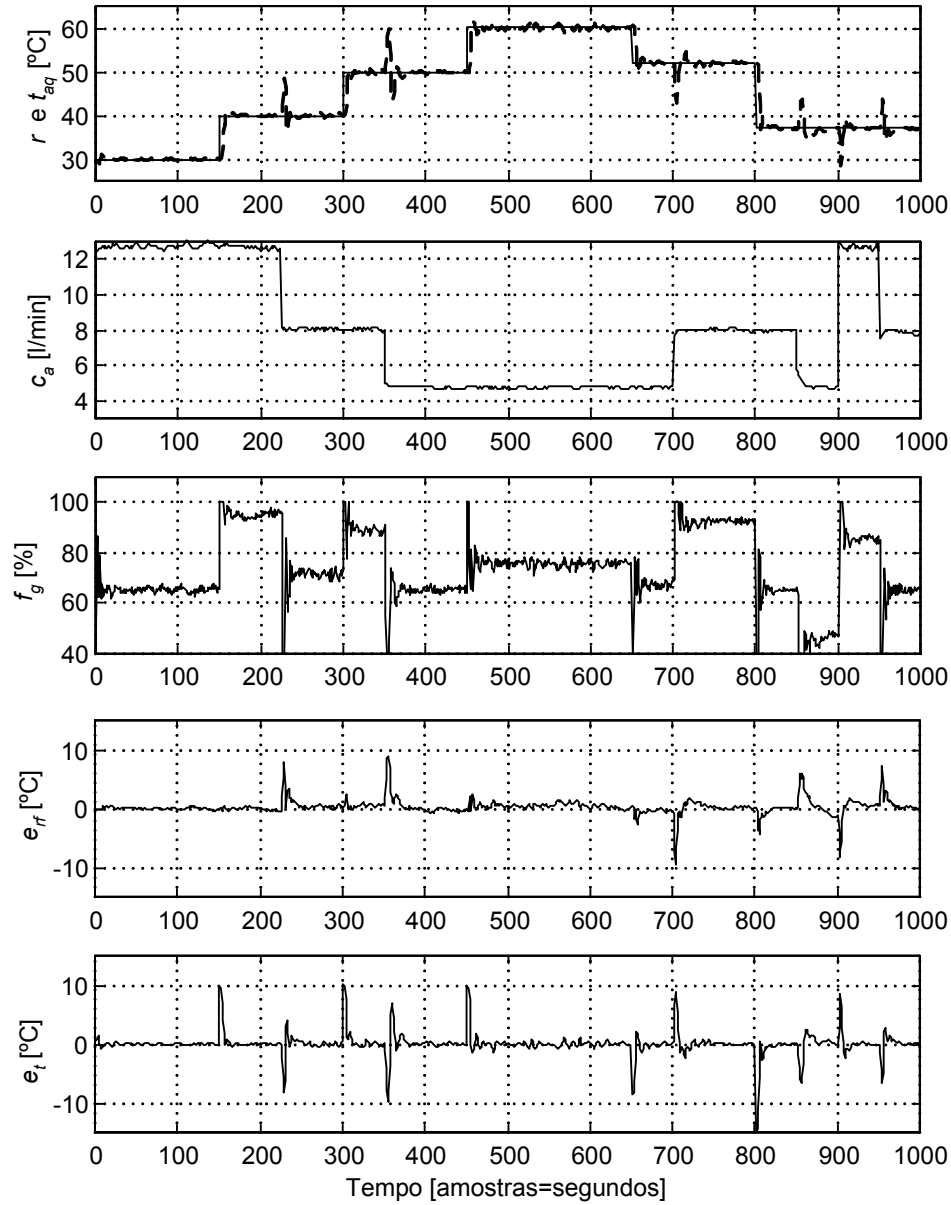


Figura 9.10: Resultados de controlo com o CSPHA.

identificado pelo modelo adaptativo.

### 9.7.3 Comentários

Os resultados apresentados na figura 9.10 obtidos pela malha de controlo Hammerstein adaptativa são satisfatórios. Estes resultados apresentam, relativamente às variações da referência, sobrelevações/subelevações que variam de 0,3 a 1,0°C obtendo sempre valores inferiores ou iguais ao valor máximo admitido. Os resultados de controlo, relativamente às variações do caudal de água e do sinal de referência, apresentam tempos de estabelecimento que variam de 5 a 14 segundos sendo sempre inferiores ao



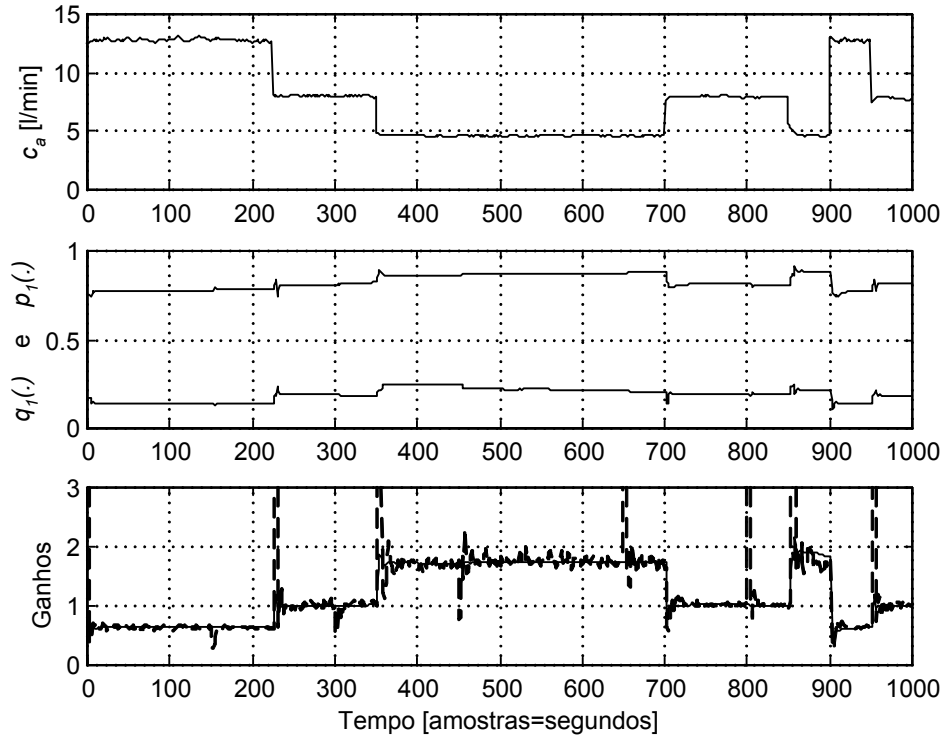


Figura 9.11: Resultados de controlo com o CSPHA (continuação).

valor máximo admitido.

Os resultados de controlo respeitam os requisitos de controlo e permitem obter um erro quadrático médio “baixo”.

Como se pode observar através da figura 9.11, perante variações de caudal, o controlador Hammerstein adaptativo necessita de algumas iterações até identificar os novos parâmetros. Os períodos de adaptação do modelo linear acabam por se reflectir negativamente no desempenho de controlo, uma vez que nestes períodos o modelo utilizado não é o correcto.

Como a identificação dos parâmetros do modelo linear dinâmico não se realiza numa só iteração, durante essas iterações os parâmetros não estão correctos. Essa incorrecção vai reflectir-se nos resultados de controlo (oscilação da temperatura da água quente em variações de caudal); contudo, os resultados são considerados satisfatórios.

A observação da evolução do sinal  $e_{rf}(k)$  (figura 9.10) permite verificar que o modelo e o esquentador respondem de forma muito semelhante. Como seria de esperar, as diferenças verificadas só ocorrem quando se verificam variações do caudal de água, uma vez que os parâmetros do modelo variam essencialmente com esta variável.

Através dos gráficos 2 e 3 da figura 9.11 pode observar-se que os parâmetros do modelo linear dinâmico variam essencialmente com o caudal de água. No entanto, nota-se alguma variação destes com as variações da temperatura da água quente. Estas variações podem ser originadas pela existência de algumas perdas de energia não modelizadas pela função não linear estática. Este problema é resolvido pela adap-

tabilidade da parte linear do modelo do esquentador, obtendo um modelo global do esquentador com um comportamento muito próximo ao comportamento do sistema real.

## 9.8 Controlador Smith Preditivo Híbrido

### 9.8.1 Malha de Controlo

A figura 9.12 apresenta o diagrama de blocos do controlador Smith preditivo híbrido série (CSPHS) polinomial. Neste caso, o modelo do esquentador é um modelo híbrido série baseado no modelo físico do esquentador. Este modelo é um sistema de 1ª ordem em que os parâmetros variam de forma não linear com o caudal de água. O modelo do esquentador é um sistema de 1ª ordem variável cujos parâmetros são calculados a cada iteração.

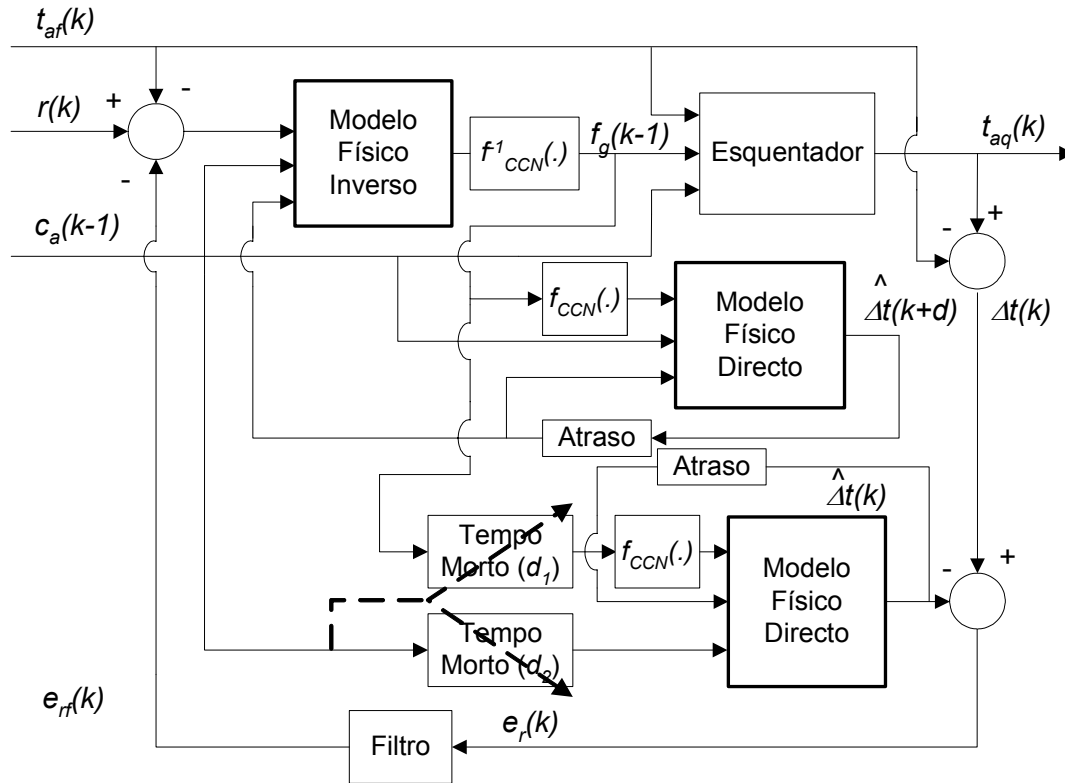


Figura 9.12: Controlador Smith preditivo híbrido série polinomial.

Apesar do modelo híbrido série polinomial ser a série de uma função não linear estática  $f_{CCN}(\cdot)$  com um modelo linear dinâmico variável  $g(\cdot)$  este esquema de controlo pode ser simplificado. Esta simplificação é efectuada da mesma forma e nas mesmas condições referidas para a simplificação do controlador Smith preditivo Hammerstein adaptativo. Neste caso, o respectivo diagrama de blocos do controlador Smith preditivo híbrido série simplificado está ilustrado na figura 9.13.



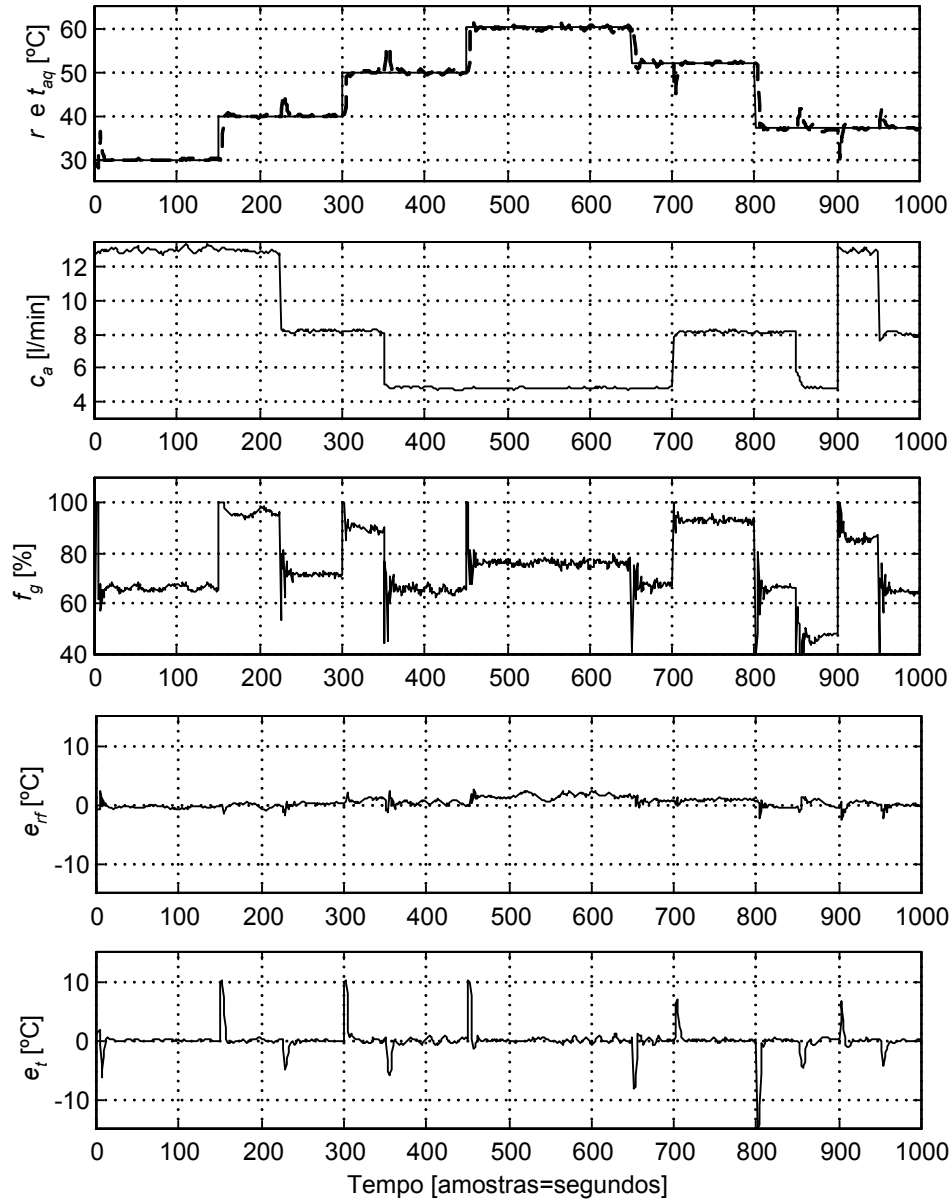


Figura 9.14: Resultados de controlo com o CSPHS.

Os resultados de controlo obtidos com o controlador Smith preditivo híbrido série apresentam uma raiz quadrada do erro quadrático médio de  $1,73^{\circ}\text{C}$ . O tempo de execução deste algoritmo de controlo é de  $1,84 \times 10^{-3}$  segundos.

Na figura 9.15, de cima para baixo, são apresentados, no gráfico 1 - o caudal da água, no gráfico 2 - os parâmetros  $K_1(k)$  e  $K_2(k)$  do modelo linear de 1ª ordem calculados a cada iteração e no gráfico 3 - os ganhos do modelo linear de 1ª ordem calculados com base nos parâmetros  $K_1(k)$  e  $K_2(k)$  (linha contínua) e com base nos sinais do sistema real linearizado (linha tracejada).

Como já tinha sido observado no modelo/controlador Hammerstein, os parâmetros do modelo linear variam com o caudal de água, assim como o respectivo ganho do

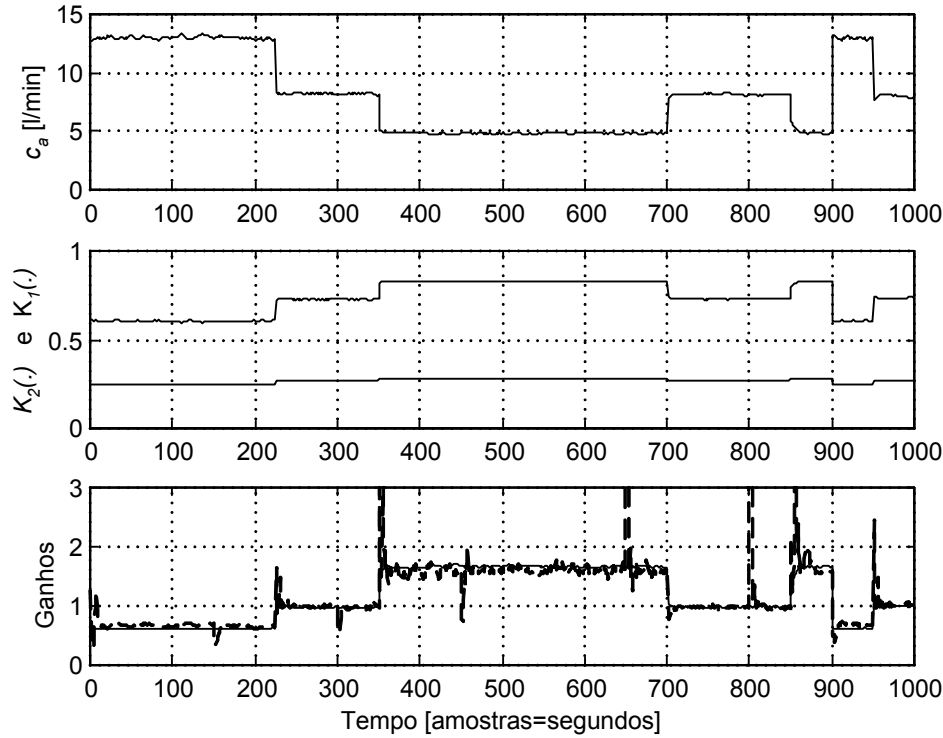


Figura 9.15: Resultados de controlo com o CSPHS (continuação).

sistema.

### 9.8.3 Comentários

Os resultados de controlo obtidos com o controlador híbrido série apresentam, relativamente às variações da temperatura de referência, sobrelevações /subelevações inferiores ao valor máximo admitido variando de 0,1 a 0,9°C. Relativamente às variações do caudal de água e da referência, os resultados de controlo apresentam tempos de estabelecimento obtidos variam de 5 a 12 segundos sendo sempre inferiores ao valor máximo admitido.

No controlador híbrido série, a mudança de parâmetros do modelo físico é efectuada a cada iteração de forma precisa e bem definida através da função transferência do sistema, obtendo o modelo correcto a cada iteração. Deste modo, o controlador não precisa de corrigir qualquer diferença entre o modelo e o sistema real, respondendo de forma satisfatória.

A observação da evolução do sinal  $e_{rf}(k)$  permite verificar que o modelo e o esquentador respondem de forma muito semelhante, uma vez que só se verificam pequenas diferenças nas variações do sinal de referência e do caudal. Verifica-se ainda um pequeno erro de modelização em estado estacionário entre o modelo e o sistema que varia com a temperatura da água quente.

O erro de modelização referido no parágrafo anterior pode ser originado pela existên-

cia de algumas perdas de energia não modelizadas pela função desconhecida  $f_{CCN}(\cdot)$ . Este problema pode ser resolvido introduzindo, na função desconhecida, mais regressores, como referido no capítulo 8, obtendo um modelo global do esquentador mais próximo do sistema real. Ao nível dos resultados de controlo, esta pequena diferença entre o modelo híbrido e o sistema real não é significativa, obtendo resultados semelhantes.

Dos gráficos 2 e 3 da figura 9.15 pode observar-se que os parâmetros do modelo físico variam com o caudal de água, como já se tinha verificado no modelo Hammerstein.

## 9.9 Comparação de Resultados

A tabela 9.1 apresenta as raízes quadradas dos erros quadráticos médios ( $RMSEs$ ) que surgem da aplicação dos perfis de teste no controlo do esquentador, as sobreelevações/subelevações relativas às variações do sinal de referência, os tempos de estabelecimento relativos às variações do caudal de água e do sinal de referência e, finalmente, os tempos de execução para as quatro malhas de controlo.

Malha de Controlo	$RMSEs$	Sobreelevações /subelevações	Tempos de estabelecimento	Tempos de Execução
	[°C]	[°C]	[segundos]	[mili-segundos]
CSPL	2,28	3 a 10	14 a 40	1,50
CSPND	1,72	0,1 a 0,9	4 a 10	3,62
CSPHA	1,90	0,3 a 1,0	5 a 14	2,75
CSPHS	1,73	0,1 a 0,9	5 a 12	1,84

Tabela 9.1: Desempenhos, sobreelevações/subelevações, tempos de estabelecimento e tempos de execução dos quatro controladores.

Os valores de  $RMSE$  apresentados na tabela possuem uma média “relativamente elevada”, uma vez que possuem uma componente que surge do erro entre o sinal de referência e a temperatura da água quente durante o tempo morto relativo à entrada fluxo de gás<sup>1</sup>. Esta componente é idêntica para todos os controladores.

Através da tabela 9.1 pode-se observar que o erro quadrático obtido pelo controlador baseado no modelo linear é muito maior do que os erros obtidos com os restantes controladores. Os resultados da aplicação deste controlador apresentam um erro quadrático médio elevado e não cumpre os requisitos de controlo (sobreelevações/subelevações relativas às variações do sinal de referência abaixo de 1°C e tempos de estabelecimento abaixo de 15 segundos).

Os controladores baseados nos modelos cinzentos não lineares obtêm erros quadráticos baixos e semelhantes e permitem cumprir os requisitos de controlo definidos.

O controlador Smith preditivo Hammerstein adaptativo obtém erros quadráticos um pouco maiores, uma vez que a adaptação dos parâmetros não é realizada numa

<sup>1</sup>Tome-se por exemplo, a evolução do sinal  $e_t(k)$  ilustrado na figura 9.14. Através deste sinal é perfeitamente visível a presença de um erro significativo quando ocorrem transições do sinal de referência.

única iteração. Nestes períodos de adaptação os parâmetros do modelo não são os correctos. Este facto reflecte-se nos resultados de controlo, contudo, o desempenho deste controlador é considerado satisfatório.

Pode ainda observar-se pela tabela 9.1 que os tempos de execução variam de controlador para controlador. Estas diferenças decorrem das diferentes cargas computacionais das várias malhas de controlo que estão intimamente ligadas ao nível de interpretabilidade e complexidade dos modelos directos e inversos utilizados. Os modelos com maior nível de interpretabilidade e menor complexidade proporcionam malhas de controlo com menor complexidade computacional (tempos de execução mais baixos).

Embora este trabalho não tivesse como prioridade a execução dos algoritmos de controlo desenvolvidos no próprio microcontrolador, estes foram implementados e executados no microcontrolador 89C51RD observando os tempos de execução e espaço em memória RAM e ROM necessários para cada algoritmo. Este estudo confirma que os tempos de execução aumentam com o aumento da complexidade dos algoritmos de controlo/modelo utilizados. Quando executados num microcontrolador as diferenças de complexidade ainda são mais visíveis devido ao seu baixo poder de cálculo. Este estudo que pode ser encontrado no Apêndice D - Sistema de Controlo do Esquentador - *Software*.

## 9.10 Conclusões

Este capítulo apresentou as várias malhas de controlo propostas para a automatização do esquentador e os vários resultados alcançados. Nas secções 9.2 e 9.3 foram definidas as condições de teste e apresentados os critérios que levaram à escolha da malha de controlo baseada em modelo. De seguida foi descrita a malha de controlo Smith preditiva genérica a aplicar na automatização do esquentador. Nas secções 9.5 até à 9.8 foram definidas as quatro malhas de controlo Smith preditivo: linear, neuro-difuso, Hammerstein adaptativo e híbrido série e apresentados os resultados de controlo e respectivos tempos de execução. Por fim, na secção 9.9, foram comparados e avaliados os resultados obtidos pelos quatro controladores apresentados.

O controlador Smith preditivo linear apresenta, como era de esperar, um valor para o erro muito elevado e não cumpre os requisitos de controlo exigidos. Este controlador foi implementado, essencialmente, para fins comparativos.

Os controladores Smith preditivos neuro-difuso, Hammerstein adaptativo polinomial e híbrido série polinomial, apresentam valores de erros baixos e semelhantes, cumprindo os requisitos de controlo exigidos no controlo do esquentador.

A complexidade matemática das malhas de controlo propostas depende fortemente da complexidade matemática dos respectivos modelos. A relação de complexidade entre os três controladores segue a relação de complexidade entre os três modelos cinzentos apresentada no capítulo anterior. Dos três controladores não lineares propostos neste capítulo, o controlador Smith preditivo híbrido série é aquele que conduz a algoritmos de menor complexidade.





# Capítulo 10

## Conclusões e Trabalhos Futuros

### 10.1 Conclusões

O trabalho apresentado nesta dissertação mostra que a utilização de modelos com um elevado nível de interpretabilidade propicia a construção de estruturas com baixa complexidade matemática e potencia a simplicidade das malhas de controlo.

A validação desta tese passou pela exploração vários métodos de modelização cinzenta, aplicados na modelização e controlo de um esquentador doméstico. Estes métodos estudam as diversas formas de combinar os vários tipos de conhecimentos na construção de modelos com um determinado grau de interpretabilidade (modelos do tipo caixa cinzenta ou cinzentos). Neste sentido, foram abordados os métodos de modelização neuro-difusa, orientada por blocos e híbrida que levaram à construção e identificação de três modelos diferentes do esquentador.

Com base nos modelos referidos foram propostas três malhas de controlo baseadas na estrutura de controlo Smith preditiva. Finalmente, foram comparados os resultados de controlo e a complexidade matemática dos respectivos algoritmos, uma vez que estes se destinam a ser executados por sistemas embutidos de baixo poder de cálculo. De referir que a avaliação das várias soluções de controlo apresentadas, além de simuladas antecipadamente num computador pessoal, foram realizadas com dados reais provenientes do esquentador.

Esta dissertação apresenta diversas contribuições, propondo diferentes abordagens de modelização e diversos algoritmos de controlo a aplicar num esquentador doméstico.

As contribuições deste trabalho são:

- Exploração de várias metodologias de combinação de conhecimentos *a priori* para a construção de modelos (modelização cinzenta);
- Definição e identificação do modelo neuro-difuso, dos modelos Hammerstein adaptativo polinomial e neuro-difuso e dos modelos híbridos série polinomial e neuro-difuso para um esquentador doméstico;
- Proposta de definição da equação de balanço de energias para um esquentador doméstico;

- Proposta de simplificação da malha de controlo Smith preditivo de forma a linearizar o sistema, relativamente a uma dada variável de entrada;
- Apresentação da malha de controlo Smith preditivo para sistemas com múltiplas entradas com diferentes tempos mortos;
- Aplicação dos vários modelos cinzentos no controlo automático de um esquentador doméstico, obtendo controladores adaptativos e não adaptativos relativamente simples.

Esta dissertação disponibiliza uma gama diversificada de soluções de controlo automático para o esquentador com semelhantes níveis de desempenho mas com diferentes níveis de complexidade matemática e interpretabilidade.

Os métodos de modelização e controlo apresentados levaram à resolução do problema proposto inicialmente e serviram de suporte e demonstração da tese apresentada nesta dissertação.

## Modelização do Tipo Caixa Cinzenta

Esta dissertação desenvolve diversas formas de combinar vários tipos de conhecimentos para a construção de modelos (modelização cinzenta), desde modelos que utilizam um simples conjunto de dados de entrada e saída (modelização cinzenta escura), passando por modelos construídos com base em conhecimentos empíricos e/ou experimentais (modelização cinzenta), até modelos baseados nas equações diferenciais fenomenológicas (modelização cinzenta clara).

Os diferentes tipos de conhecimentos permitem construir modelos com diferentes níveis de interpretabilidade, de complexidade matemática e de capacidade de generalização. Esta é uma área muito vasta onde muito trabalho pode ainda ser desenvolvido.

## Modelos Neuro-Difuso, Hammerstein Adaptativo e Híbrido Série

Da exploração destas metodologias de modelização foram definidos e identificados três tipos de modelos para o esquentador: neuro-difuso, Hammerstein adaptativo e híbrido série, estes dois últimos com duas variantes cada um.

Todos os modelos desenvolvidos utilizaram conhecimentos empíricos e/ou experimentais sobre o esquentador que permitiram a redução do número de entradas do sistema e a definição dos regressores.

Os erros de modelização obtidos pelos três modelos são idênticos; no entanto, a sua complexidade matemática e interpretabilidade diferem de modelo para modelo. O modelo neuro-difuso define a sua estrutura e ajusta os seus parâmetros internos, fundamentalmente, à custa de um conjunto de dados de entrada e saída. Dos vários modelos propostos este é o que apresenta uma estrutura com maior complexidade e menor interpretabilidade.

A estrutura do modelo Hammerstein é composta por dois blocos ligados em série: uma função não linear estática e um modelo linear dinâmico. Estes blocos são definidos

com base em conhecimentos físicos e/ou experimentais e ajustados com base num conjunto de dados de entrada e saída. Estes modelos possuem, relativamente aos restantes modelos propostos, uma complexidade e interpretabilidade intermédias.

O modelo híbrido define a sua estrutura com base em conhecimentos físicos do esquentador (equações de balanço de energia) sendo constituído por dois blocos: um modelo físico e um modelo do tipo caixa cinzenta ou negra. A validação do modelo físico e ajuste dos parâmetros internos da função desconhecida do tipo caixa cinzenta ou negra são efectuados com base num conjunto de dados de entrada e saída. Este é o modelo que possui menor complexidade matemática, maior nível de interpretabilidade e maior capacidade de generalização.

## **Equação de Balanço de Energias do Esquentador**

A construção do modelo híbrido série do esquentador exigiu a proposta de uma definição da equação de balanço de energias deste termodoméstico de forma a obter o respectivo modelo físico. Este modelo é definido, fundamentalmente, através das equações físicas do sistema. Porém, existe uma componente deste modelo que é desconhecida, sendo modelizada através de uma função do tipo caixa cinzenta ou negra.

A equação de balanço de energias do esquentador permitiu, conjuntamente com o modelo discreto da função desconhecida, a definição da função transferência do esquentador no domínio discreto. Constatou-se que a função transferência do esquentador obtida desta forma conduz a uma boa aproximação do comportamento do sistema.

## **Linearização do Sistema Relativamente a uma Dada Variável de Entrada**

O facto dos modelos desenvolvidos possuírem um razoável grau de interpretabilidade e de serem constituídos por blocos, levaram à proposta de simplificação da malha de controlo Smith preditiva, linearizando o sistema relativamente ao fluxo de gás. Esta linearização só é aplicada aos controladores Smith preditivos Hammerstein e híbridos, uma vez que só estes controladores é que possuem modelos constituídos por dois blocos em série separáveis.

A linearização do sistema relativamente a uma dada variável de entrada simplifica, em grande medida, a complexidade dos respectivos controladores.

## **Malha de Controlo Smith Preditiva para Sistemas MISO**

O controlo Smith preditivo de sistemas MISO com diferentes tempos mortos requer, conforme o capítulo 9, de alguns cuidados importantes. Este trabalho apresenta os cuidados a ter em conta na configuração da malha de controlo Smith preditiva, propondo a utilização de dois modelos directos: um com tempos mortos e outro sem os referidos tempos. Este procedimento é fundamental para que a malha de controlo Smith preditiva de sistemas MISO, com diferentes tempos mortos, obtenha resultados de controlo satisfatórios.

## Resultados de Controle

A aplicação dos vários modelos cinzentos no controlo Smith preditivo do esquentador permitiu obter controladores adaptativos e não adaptativos. Genericamente, os controladores não adaptativos desenvolvidos para o esquentador (controlador Smith preditivo neuro-difuso e híbrido série) possuem modelos com parâmetros fixos implicando algoritmos de controlo de média/baixa complexidade matemática. O controlador adaptativo desenvolvido para o esquentador (controlador Smith preditivo Hammerstein) possui um modelo linear muito simples mas, com a utilização de um algoritmo recursivo de adaptação de parâmetros em linha, a sua complexidade aumenta consideravelmente.

Os resultados de controlo obtidos pelos três controladores propostos respeitam largamente as especificações de desempenho exigidas pelos vários fabricantes de esquentadores.

## Modelo/Controlador a Implementar no *Hardware*

O *hardware* definido para o controlo do sistema é baseado num microcontrolador de baixo poder de cálculo, como os desempenhos dos vários controladores são semelhantes, a escolha do controlador a implementar deverá basear-se no algoritmo que possuir menor complexidade matemática (menor tempo de execução). Dos três controladores propostos o controlador Smith preditivo baseado no modelo híbrido série polinomial (CSPHS) é o mais adequado para implementar no *hardware* existente.

## 10.2 Trabalhos Futuros

Os trabalhos futuros propostos surgem de problemas que foram identificados no decorrer deste trabalho e que ficaram em aberto, merecendo um estudo mais aprofundado.

Dos problemas identificados na modelização e com o controlo automático do esquentador que merecem mais alguma atenção, destacam-se os seguintes:

- Modelização Física do Esquentador;
- Modelização Neuro-Difusa: Interpretabilidade *versus* Complexidade;
- Modelização/Controlo com Tempo Morto Fraccionário;
- Modelização/Controlo Adaptativo do Esquentador.

Do conhecimento fenomenológico e experimental do sistema, surge a ideia de estudar e testar a modelização/controlo do esquentador, através da seguinte técnica:

- Modelização/Controlo do Esquentador com Período de Amostragem Variável.

## Modelização Física do Esquentador

A definição da equação de balanço de energia do sistema deveria ser completa permitindo uma modelização física completa do esquentador. A definição da potência térmica fornecida pelo fluxo de gás e da potência térmica libertada para o exterior deveria ser estudada mais profundamente, uma vez que, quanto maior for o seu conhecimento, menores serão as componentes desconhecidas e melhor será o modelo obtido.

Embora a equação física correspondente à potência térmica fornecida por um determinado fluxo de gás seja possível de definir, a equação correspondente à potência térmica libertada para o exterior é difícil de obter já que possui parâmetros que variam de forma desconhecida. A definição destas equações pode simplificar o modelo final obtendo um modelo do tipo caixa branca para o esquentador.

## Modelização Neuro-Difusa: Interpretabilidade *versus* Complexidade

Os modelos neuro-difusos podem possuir diferentes níveis de interpretabilidade e complexidade matemática. Os modelos deste tipo apresentados neste trabalho possuem níveis de interpretabilidade relativamente baixos e níveis de complexidade relativamente elevados. Todavia, impondo algumas restrições às respectivas estruturas é possível aumentar os seus níveis de interpretabilidade, propiciando a redução da respectiva complexidade.

Já existem várias abordagens a este problema, como as apresentadas por Paiva [Pai99] e Babuska [Bab98] que utilizam técnicas de fusão de funções de pertinência para aumentar a interpretabilidade dos respectivos modelos reduzindo desta forma a sua complexidade. Babuska [Bab98] apresenta outra técnica para a resolução deste problema que consiste na imposição de funções lineares locais interpretáveis nos consequentes das regras neuro-difusas. O objectivo é explorar técnicas que permitem o aumento da interpretabilidade dos modelos neuro-difusos potenciando a diminuição da complexidade matemática destas estruturas.

## Modelização/Controlo com Tempo Morto Fraccionário

A modelização dos tempos mortos fraccionários existentes no esquentador não foi efectuada. Será interessante verificar em que medida é que a modelização dos tempos mortos fraccionários melhoraria os vários modelos apresentados.

Sabendo que a existência de tempos mortos fraccionários não modelizados implica uma degradação dos resultados de controlo, seria interessante estudar essa degradação em função do valor do tempo morto fraccionário.

A redução da degradação de desempenho provocada pela não consideração da existência dos tempos mortos fraccionários pode passar, simplesmente, pela diminuição “controlada” do período de amostragem permitindo reduzir o tempo morto fraccionário. Desta forma a relação entre o tempo morto fraccionário e a dinâmica do sistema diminui, reduzindo a perturbação introduzida por um tempo morto fraccionário ignorado.

A modelização dos tempos mortos inteiros e fraccionários pode ser efectuada utilizando vários regressores do fluxo de gás, incluindo deste modo o tempo morto inteiro e fraccionário no modelo do sistema. Existem outras técnicas para a modelização dos tempos mortos inteiros e fraccionários como, a predição recursiva e não recursiva neuronal do tempo morto em sistemas lineares e não lineares [TC99].

## **Modelização/Controlo Adaptativo do Esquentador**

Os esquentadores domésticos podem funcionar com vários tipos de gás e possuir várias potências estáticas máximas. Neste sentido, será relevante estudar as vantagens/desvantagens da utilização de modelos neuro-difusos e híbridos (parte desconhecida) adaptativos, como é efectuado no controlador Hammerstein adaptativo proposto.

Por outro lado, como no esquentador não actuam perturbações/cargas desconhecidas (tendo conhecimento de todas as variáveis do sistema) o auto-ajuste do modelo do esquentador numa fase inicial talvez seja a solução mais indicada. Por isso, a solução passaria por existir uma fase inicial de auto-ajuste em linha que permitisse a identificação do modelo relativamente à potência estática máxima e ao tipo de gás consumido.

## **Modelização/Controlo do Esquentador com Período de Amostragem Variável**

Do conhecimento fenomenológico e experimental do sistema, surge a ideia de modelizar/controlar o esquentador utilizando um período de amostragem variável. Como o esquentador pode ser modelizado correctamente por um modelo linear de 1ª ordem, com ganho e constante de tempo variáveis com o caudal de água e de forma conhecida, a proposta sugere a variação do período de amostragem em linha de forma a construir e identificar um modelo linear de 1ª ordem com constante de tempo “fixa” e ganho variável com o caudal de água. Este período de amostragem irá variar com o caudal de água de forma a manter “fixa” a constante de tempo do esquentador relativamente ao respectivo período de amostragem. Este procedimento permite a criação de um modelo para o esquentador muito simples e de elevado nível de interpretabilidade. A imposição de um período de amostragem variável para o controlo do sistema poderá trazer alguns problemas, nomeadamente ao nível da aquisição de dados de entrada e saída e da definição dos tempos mortos variáveis do sistema.

Sobre esta ideia muito trabalho terá de ser efectuado por forma a poder comparar a qualidade do modelo construído e o desempenho dos resultados de controlo, relativamente aos resultados obtidos pelos modelos e controladores apresentados nesta dissertação.

# Bibliografia

- [AB00] J. Abonyi and R. Babuska. Local and global identification and interpretation of parameters in Takagi-Sugeno fuzzy models. *Proceedings of the 9th IEEE International Conference on Fuzzy Systems*, pages 835–840, 2000.
- [ABB<sup>+</sup>00] J. Abonyi, R. Babuska, M. Botto, F. Szeifert, and L. Nagy. Identification and control of nonlinear systems using fuzzy Hammerstein models. *Industrial Engineering and Chemistry Research*, 39(11):4302–4314, 2000.
- [Abo03] J. Abonyi. *Fuzzy Model Identification for Control*. Birkhauser, 2003.
- [Abr01] A. Abraham. *Neuro-Fuzzy Systems: State of the Art Modelling Techniques*. Connectionist Models of Neurons, Learning Processes, and Artificial Intelligence, Lecture Notes in Computer Science. Volume. 2084, Springer Verlag Germany, Jose Mira and Alberto Prieto (Eds.), Spain, pp. 269-276, 2001.
- [ABVS00] J. Abonyi, R. Babuska, H. Verrbuggen, and F. Szeifert. Incorporating à priori knowledge in fuzzy model identification. *Journal of System Science*, 31(5):657–667, 2000.
- [Aka73] H. Akaike. Information theory and an extension of the maximum likelihood principle. *2nd International Symposium on Information Theory*, pages 267–281, 1973.
- [AMM96] N. Alves, H. Macedo, and J. Maciel. Gas heater control using fuzzy logic. *Internal Report - Instituto Superior Técnico da Universidade Técnica de Lisboa Portugal*, 1996.
- [AMS02] J. Abonyi, J. Madar, and F. Szeifert. *Soft Computing in Industrial Applications - Recent Advances, Combining First Principles Models and Neural Networks for Generic Model Control in Soft Computing in Industrial Applications - Recent Advances Book*. Eds. R. Roy, M. Koppen, S. Ovaska, T. Furuhashi, F. Homann Springer Engineering Series, 2002.
- [AN00] A. Abraham and B. Nath. Hybrid intelligent systems: A review of a decade of research. Technical Report 5, School of Computing and Information Technology, Faculty of Information Technology , Monash University, Australia, Technical Report Series, pp. 1-55, 2000.

- [ARBS03] J. Abonyi, H. Roubos, R. Babuska, and F. Szeifert. *Identification of Semi-Mechanistic Models with Interpretable TS-Fuzzy Submodels by Clustering, OLS and FIS Model Reduction, Fuzzy Modeling and the Interpretability-Accuracy Trade-Off. Part I, Interpretability Issues., Chapter 10.* J. Casillas, O. Cordon, F. Herrera, L Magdalena, (Eds.), Studies in Fuzziness and Soft Computing, Physica-Verlag, 2003.
- [AW84] K. J. Ashtrom and B. Wittenmark. *Computer Controlled Systems: Theory and Design.* Prentice Hall, 1984.
- [Bab98] R. Babuska. *Fuzzy Modeling for Control.* Kluwer Academic Publishers, Boston, 1998.
- [Bai98] E. W. Bai. An optimal two-stage algorithm for Hammerstein-Wiener nonlinear systems. *Automatica*, 34(3):333–338, 1998.
- [BB99] G. Bontempi and M. Barattari. *Toolbox for Neuro-Fuzzy Identification and Data Analysis - for Use with Matlab.* Université Libre de Bruchelles, 50, av. Fraklin Roosevelt 1050 Bruchelles - Belgium, 1999.
- [Bej93] A. Bejan. *Heat Transfer.* John Wiley and Sons, 1993.
- [Bez73] J. C. Bezdek. *Fuzzy Mathematics in Pattern Classification.* PhD thesis, Applied Math. Center, Cornell University, Ithaca, 1973.
- [BFOS84] L. Breiman, J. Freidman, R. Olshen, and C. Stone. *Classification and Regression Trees.* Wadswoks, Inc., Belmont, California, 1984.
- [BH91] M. Brown and C. J. Harris. A nonlinear adaptive controller: A comparison between fuzzy logic and neurocontrol. *IMA Journal of Mathematical Control and Information*, 8(3):239–265, 1991.
- [BK92] H. R. Berenji and P. Khedkar. Learning and tuning fuzzy logic controllers through reinforcements. *IEEE Transactions on Neural Networks*, 3:724–740, 1992.
- [Bos97] K. Bossley. *Neurofuzzy Modelling Approches in Systems Identification.* PhD Thesis, Departement of Electronics and Computer Science, Falculty of Engineering and Applied Science, University of Southampton, United Kingdom, 1997.
- [BRD96] M. Boutayeb, H. Rafaralahy, and M. Darouach. A robust a recursive identification method for Hammerstein model. *Proc. Of the 13th Triennial IFAC World Congress*, 1:447–452, 1996.
- [BV97] R. Babuska and H. B. Verbruggen. *Fuzzy Set Methods for Local Modeling and Identification.* In R. Murray-Smith and T. A. Johansen, Editors *Multiple Model Approaches to Modelling and Control.* Taylor and Fracis, 1997.



- [BV03] R. Babuska and H. Verbruggen. Neuro-fuzzy methods for nonlinear system identification. *Annual Reviews in Control*, 27:73–85, 2003.
- [CAPL96] F. Cubillos, P. Alvarez, J. Pinto, and E. Lima. Hybrid-neural modelling for particulate solid drying processes. *Power Tecnology*, 87:153–583, 1996.
- [CB99] E. Camacho and C. Bordons. *Model Predictive Control*. Springer Verlag, 1999.
- [CBH<sup>+</sup>96] H. Can, H. Braake, C. Hellinga, K. Luyben, and J. Heijnen. Strategy for dynamic process modelling based on neural networks and macroscopic balances. *AIChE*, 42:3403–3418, 1996.
- [CCR94] S. Claquin, A. Carriere, and F. Rocaries. Modelling and application of adaptive control to a gas heater. *Proceedings of the Third IEEE Conference on Control Applications*, 2:927–932, 1994.
- [Chi94] S. L. Chiu. Fuzzy model identification based on cluster estimation. *Journal of Intelligent and Fuzzy Systems*, 2(3), 1994.
- [Chi02] N. Chiras. *Linear and Nonlinear Modelling of Gas Turbine Engines*. PhD thesis, University of Glamorgan, 2002.
- [CMT87] D. Clarke, C. Mothadi, and P. Tuffs. Generalized predictive control. part I the basic algorithm and part II extensions and interpretations. *Automatica*, 23(2):137–160, 1987.
- [Coe02] M. Coelho. *Modelos de Hammerstein e de Wiener : Conexões Com Modelos NARX Polinomiais e Sua Aplicação Em Identificação de Sistemas Não Lineares*. PhD thesis, MsD, PPGEE, UFMG, 2002.
- [Cor01] M. V. Correa. *Identificação Caixa-Cinza de Sistemas Não-Lineares Utilizando Representações NARMAX Racionais e Polinomiais*. PhD thesis, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil, 2001.
- [DC03] D. Dougherty and D. Cooper. Tuning guidelines of a dynamic matrix controller for integrating (non-self-regulating) processes. *Ind. Eng. Chemical Research*, 42:1739–1752, 2003.
- [DHR93] D. Driankov, H. Hellendoorn, and M. Reinfrank. *An Introduction to Fuzzy Control*. Springer-Verlag, Heidelberg, Germany, 1993.
- [DS81] N. R. Draper and H. Smith. *Applied Regression Analysis*. John Wiley, New York, 1981.
- [EJL91a] E. Eskinat, S. H. Johnson, and W. Luyben. Use of Hammerstein models in identification of nonlinear systems. *AIChE Journal*, 37(2):255–268, 1991.
- [EJL91b] E. Eskinat, S. Johnson, and W. Luyben. Use of hammerstein models in identification of nonlinear systems. *Journal of AIChE*, 37:255–268, 1991.

- [EMP86] C. Economou, M. Morari, and B. Palsson. Internal model control. v extension to nonlinear systems. *Ind. Eng. Chemical Process Res. Dev.*, 25:403–411, 1986.
- [Eyk74] J. Eykhoff. *System Identification*. John Wiley and Sons, London UK, 1974.
- [FG99] M. Figueiredo and F. Gomide. Design of fuzzy systems using neuro-fuzzy networks. *IEEE Transactions on Neural Networks*, 10(4):815–827, 1999.
- [Fra88] Arthur Fraas. *Heat Exchanger Design, Second Edition*. John Wiley and Sons, 1988.
- [GM82] C. Garcia and M. Morari. Internal model control i - a unifying review and some new results. *Ind. Eng. Chemical Process Res. Dev.*, 21:308–323, 1982.
- [HA93] X. He and H. Asada. A new method for identifying orders of input-output models for nonlinear dynamic systems. *Proc. of American Control Conference*, 1:2520–2523, 1993.
- [Hab85] H. Haber. Nonlinearity tests for dynamic processes. *In IFAC Symposium on Identification and system parameter estimation*, pages 409–413, 1985.
- [HBC92] Y. Hayashi, J. J. Buckley, and E. Czogala. Systems engineering applications of fuzzy neural networks. *International Joint Conference on Neural Networks*, 2(7-11):413 – 418, 1992.
- [HCRH94] E. Huyghe, S. Claquin, F. Rocaries, and Y. Hamam. An adaptive fuzzy controller for a non linear thermal system. *IEEE International Conference on Systems, Man, and Cybernetics Humans, Information and Technology*, 1:448–452, 1994.
- [Hen98] M. Henson. Nonlinear model predictive control: Current status and future directions. *Computers and chemical engineering*, 23:187–202, 1998.
- [HH95] E. Huyghe and Y. Hamam. Simulated annealing for fuzzy controller optimization: Principles and applications. *IEEE International Conference on Intelligent Systems for the 21st Century*, 5:4509–4514, 1995.
- [HSZG92] K. Hunt, D. Sbarbaro, R. Zbikowski, and P. Gawthrop. Neural networks for control systems - a survey. *Automatica*, 28:1083–1112, 1992.
- [HUK03] T. Hatanaka, K. Uosaki, and M. Koga. Block oriented nonlinear model identification by evolutionary computation approach. *Proceedings of 2003 IEEE Conference on Control Applications*, 1:43 – 48, 2003.
- [Hum99] Humusoft. *The Real Time Toolbox 3.0 - Used with Matlab*. Humusoft, 1999.
- [Jan92] J. S. R. Jang. *Neuro-Fuzzy Modelling: Architectures, Analysis and Applications*. PhD Thesis, University of California, Berkley, 1992.

- [Jan93] J. S. R. Jang. ANFIS: Adaptive network based fuzzy inference systems. *IEEE Transactions on Systems, Man and Cybernetics*, 23(3):665–685, 1993.
- [Jan96] J. Jang. Input selection for ANFIS learning. *Proc. IEEE International Conference on Fuzzy Systems*, 2:1493–1499, 1996.
- [JH95] S. Jorgensen and K. Hangos. Grey modelling for control: Qualitative models as a unifying framework. *International Journal of adaptive control and signal processing*, 9:547–562, 1995.
- [JL98] F. C. Juang and T. Chin Lin. An on-line self constructing neural fuzzy inference network and its applications. *IEEE Transactions on Fuzzy Systems*, 6:12–32, 1998.
- [JS93] J. S. R. Jang and C-T. Sun. Functional equivalence between radial basis functions networks and fuzzy inference systems. *IEEE Transactions on Neural Networks*, 4(1):156–159, 1993.
- [JSM97] J. S. Jang, C. T. Sun, and E. Mizutani. *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. Matlab Curriculum Series - Prentice Hall, 1997.
- [Kar76] W. Karplus. *The Spectrum of Mathematical Modelling and System Simulation. Simulation and Systems*. North Holland Publishing Company, 1976.
- [KAWC95] A. Kalafatis, N. Arifin, L. Wang, and W. Cluett. A new approach to the identification of pH processes based on the Wiener model. *Chemical Eng. Science*, 50:3693–3701, 1995.
- [KFT03] S. Kim, P. Fleming, and H. Thompson. Neuro-fuzzy modeling approach to gas turbine dynamic system. *IFAC International Conference on Intelligent Control Systems and Signal Processing*, 2:532–537, 2003.
- [KK82] P. Krishnaiah and L. Kanal. *Classification, Pattern Recognition, and Reduction of Dimensionality*. Handbook of statistics (Volume 2), North-Holland, Amsterdam, 1982.
- [KL96] T. Kavli and G.T. Lines. A unifying framework for mechanistic: Fuzzy and data driven modelling. *Proceedings of EUFIT'96, Aachen, Germany, 2-5 Set*, 1996.
- [KR76] R. L. Kashyap and A. R. Rao. *Dynamic Stochastic Models from Empirical Data*. Academic Press, New York, 1976.
- [KS99] N. Kasabov and Qun Song. *Dynamic Evolving Fuzzy Neural Networks with 'm-Out-of-N' Activation Nodes for On-Line Adaptive Systems*. Technical Report TR99/04, Departement of Information Science, University of Otago, 1999.

- [LB85] I. Leontaritis and S. Billings. Input-output parametric models for non-linear systems part II: Deterministic non-linear systems. *Int. Journal Control*, 41(2):329–344, 1985.
- [LC97] J. Leski and E. Czogala. A new artificial neural network based fuzzy inference system with moving consequents in if-then rules. *BUSEFAL*, 71:72–81, 1997.
- [LC99] J. Leski and E. Czogala. A new artificial neural network based fuzzy inference system with moving consequents in if-then rules and its applications. *Fuzzy Sets and Fuzzy Logic*, 108:289–297, 1999.
- [Lev44] K. Levenberg. A method for the solution of certain problems in least squares. *Quart. Appl. Math*, 2:164–168, 1944.
- [Lin96] P. Lindskog. *Methods, Algorithms and Tools for System Identification Based on Prior Knowledge*. Department of Electrical Engineering - Linköping University Sweden, 1996.
- [Lit02] Pascal F. Van Lith. *Hybrid Fuzzy-First Principles Modeling*. PhD. Dissertation University Twente Press, 2002.
- [Lju87] L. Ljung. *System Identification: Theory For the User*. Prentice-Hall Information and System Sciences Series, 1987.
- [Lju91] L. Ljung. System identification toolbox - for use with MATLAB, user's guide. 1991.
- [Lju00] A. Hagenblad Ljung. Maximum likelihood estimation of Wiener models. *Proceedings of the 39th IEEE Conference on Decision and Control*, 3:2417–2423, 2000.
- [LL91] T. C. Lin and C. S. Lee. Neural network based fuzzy logic control and decision system. *IEEE Transactions on Computers*, 40(12):1320–1336, 1991.
- [Luy90] L. Luyben. Process modeling, simulation, and control of chemical engineers. 2nd ed. McGraw-Hill, New York (USA), 1990.
- [Mad02] J. Madar. *Thesis: Application of Hybrid and Grey-Box Modelling in the Process Engineering*. Department of Process Engineering University of Veszprém, Hungary, 2002.
- [Man76] E. H. Mandani. Advances in linguistic synthesis of fuzzy controllers. *International Journal of Man Machine Studies*, 8:669–678, 1976.
- [MAP97] P. H. Menold, F. Allgower, and R. K. Pearson. On simple representations of distillation dynamics. *Proceedings ECCE-1, Florence, Italy*, 1997.
- [Mar63] D. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *SIAM J. Appl. Math*, 11:431–441, 1963.

- [Mat99] MathWorks. *MATLAB Version 5.3 - The Language of Technical Computing*. The MathWorks Inc., January 1999.
- [Mat01] Mathworks. *Fuzzy Logic Toolbox 2.0 for Use with MATLAB. User's Guide, Version*. The Mathworks, Inc., 2001.
- [MF96] A. Mota and J. Fonseca. Sistema de controlo de um esquentador de uso doméstico - controlador fuzzy. *Relatório Interno (Confidencial) - Departamento de Electrónica e de Telecomunicações Universidade de Aveiro*, 1996.
- [MO95] J. Moscinski and Z. Ogonowski. *Advanced Control with Matlab and Simulink*. Ellis Horwood, 1995.
- [Mot97] A. Mota. *Identificação de Sistemas*. DETUA, 1997.
- [Nau94] D. Nauck. A fuzzy perceptron as a generic model for neuro-fuzzy approaches. *Proc. Fuzzy-Systems, 2nd GI-Workshop, Munich*, 1994.
- [NFBS00] O. Nelles, A. Fink, R. Babuska, and M. Setnes. Comparison of two construction algorithms for Takagi-Sugeno fuzzy models. *International journal of applied mathematics and computer science*, 10(4):835–855, 2000.
- [NG66] K. S. Naredra and P. G. Gallman. An interactive method for the identification of nonlinear systems using Hammerstein model. *IEEE Transactions on Automatic control*, 11(3):546–550, 1966.
- [NH04] C. Ng and M. Hussain. Hybrid neural network-priori knowledge model in temperature control of a semi-batch polymerization process. *Chemical Engineering and Processing*, 43:559–570, 2004.
- [NK97] D. Nauck and R. Kurse. Neuro-fuzzy systems for function approximation. *4th International Workshop Fuzzy-Neuro Systems*, 1997.
- [NKK97] D. Nauck, F. Klawonn, and R. Kruse. *Foundations of Neuro-Fuzzy Systems*. John Wiley and Sons, 1997.
- [Nor86] J. P. Norton. *An Introduction to Identification*. Academic Press, 1986.
- [NPR96] S. J. Norquary, A. Palazuglo, and J. A. Romagnoli. Nonlinear model predictive control of PH neutralization using Wiener models. *Proceedings of IFAC World Conference, San Francisco, USA*, 1996.
- [Oli98] R. Oliveira. *Supervision, Control and Optimization of Biotechnological Processes Based on Hybrid Models*. PhD. Dissertation University of Halle, 1998.
- [Oli04] R. Oliveira. Combining first principles modelling and artificial neural networks: A general framework. *Computers and Chemical Engineering*, 28 (5):755–766, 2004.

- [PA88] B. Pradeep and R. H. Ash. *Computer Process Control with Advanced Control Applications*. ISA, 1988.
- [Pai99] R. Paiva. *Identificação Neuro-Difusa - Aspectos de Interpretabilidade*. M.Sc. Thesis, Departamento de informática da Universidade de Coimbra, Portugal, 1999.
- [Ped96] W. Pedricz. Conditional fuzzy c-means. *Pattern Recognition*, 17:625–631, 1996.
- [PP98] M. Pottman and R. K. Pearson. Block-oriented narmax models with output multiplicities. *AIChE Journal*, 44:131–140, 1998.
- [PP00] R. K. Pearson and M. Pottmann. Gray-box identification of block-oriented nonlinear models. *Journal of process control*, 10:301–315, 2000.
- [PU92a] D. Psychogios and L. Ungar. A hybrid neural network - first principle approach using artificial neural networks. *Ind. Eng. Chemistry Research*, 30:2564–2580, 1992.
- [PU92b] D. C. Psychogios and L. H. Ungar. A hybrid neural network-first principles approach to process modelling. *AIChE*, 38(10):1499–1511, 1992.
- [RBQG97] Z. Rongming, T. Bian, W. Qianntu, and D. Guanzhong. Application of fuzzy logic in home appliance: Gas heater controller design. *IEEE International Conference on Intelligent Processing Systems*, 1:373–376, 1997.
- [RKS<sup>+</sup>99] J. Roubos, P. Krabben, M. Setnes, R. Babuska, J. Heijnen, and H. Verbruggen. Hybrid model development for fed-batch bioprocesses: Combining physical equations with the metabolic network and black box kinetics. *6th Workshop on Fuzzy Systems Brunel University (UK)*, 1:231–239, 1999.
- [SB04] R. Sindelar and R. Babuska. Input selection for nonlinear regression models. *IEEE Transactions on fuzzy systems*, 12(5):688–696, 2004.
- [SEM89] D. Seborg, E. Edgar, and D. Mellichamp. *Em Sistemas Com Tempo Morto oProcess Dynamics and Control, 9th Edition*. John Wiley and Sons Inc., 1989.
- [SL97] R. Simutis and A. Lubbert. Exploratory analysis of bioprocesses using artificial neural network-based methods. *Biothechnology Process*, 13:479–487, 1997.
- [SMR98] N. Silva, H. Macedo, and A. Rosa. Evolutionary fuzzy neural networks automatic design of rule based controllers of nonlinear delayed systems. *IEEE International Conference on Fuzzy Systems*, 2:1271–1276, 1998.
- [Soh98] B. Sohlberg. *Supervision and Control for Industrial Processes - Using Grey-Box Model, Predictive Control and Fault Detection Methods*. Springer-Verlag, 1998.

- [Sor94] O. Sorensen. *Neural Networks in Control Applications*. PhD thesis, Department of Control Engineering, Institute of Electronic Systems, Aalborg University, Denmark, 1994.
- [Sou03] J. Sousa. *Fuzzy Decision Making in Modeling and Control*. World Scientific Publishing Company, Incorporated, 2003.
- [SS89] T. Soderstrom and P. Stoica. *System Identification*. Prentice Hall, 1989.
- [SSD<sup>+</sup>94] J. Schubert, R. Simutis, M. Dors, I. Havlik, and A. Lubbert. Bioprocess optimization and control: Application of hybrid modelling. *Journal of biotechnology*, 35:51–68, 1994.
- [Sto81] P. Stoica. On the convergence of an interactive algorithm used for Hammerstein system identification. *IEEE Transactions Automatic Control*, 26:967–969, 1981.
- [STV93] S. Sulzberger, N. Tschichold, and S. Vestli. FUN: Optimization of fuzzy rule based systems using neural networks. *Proceedings of IEEE Conference on Neural Networks, San Francisco*, 1:312–316., 1993.
- [TC99] Y. Tan and A. Van Cauwenberghe. Neural-network-based d-step-ahead predictors for nonlinear systems with time delay. *Engineering Applications of Artificial Intelligence*, 12(1):21–35, 1999.
- [TJW96] A. Tsen, S. Jang, and D. Wong. Predictive control of quality in batch polymerization using hybrid ANN models. *Journal of the American Institute of Chemical Engineers*, 42(2):455–465, 1996.
- [TK94a] M. Thompson and M. Kramer. Modelling chemical processes using priori knowledge and neural networks. *AIChE Journal*, 40(8):1328–1340, 1994.
- [TK94b] M. Tompson and M. Kramer. Modelling chemical processes using priori knowledge and neural networks. *AIChE*, 40(8):1328–1340, 1994.
- [TKB92] M. Thompson, M. Kramer, and P. Bhagat. Embedding theoretical model in neural networks. *ACC/WA*, 14:475–486, 1992.
- [TOA96] S. Tano, T. Oyama, and T. Arnould. Deep combination of fuzzy inference and neural network in fuzzy inference. *Fuzzy Sets and Systems*, 82(2):151–160, 1996.
- [TR96] A. Tholudur and W. Ramirez. Optimization of fed-batch bioreactors using neural network parameter function models. *Biotechnology Process*, 12:302–309, 1996.
- [TS85] T. Takagi and M. Sugeno. Fuzzy identification of systems and its application to modeling and control. *IEEE Transactions on Systems, Man and Cybernetics*, 15(1):116–132, 1985.

- [TZM01] Y. Tian, J. Zhang, and A. Morris. Modelling and optimal control of a batch polymerisation reactor using a hybrid stacked recurrent neural network model. *Industrial and Engineering Chemistry Research*, 40(21):4525–4535, 2001.
- [VDM04a] J. Vieira, F. Dias, and A. Mota. Artificial neural networks and neuro-fuzzy systems in model and control: A comparative study. *International Journal ELSEVIER, Engineering Applications of Artificial Intelligence*, 17(3):265–273, 2004.
- [VDM04b] J. Vieira, F. Dias, and A. Mota. Neuro-fuzzy systems: A survey. *WSEAS Transactions on Systems*, 2(3):414–419, 2004.
- [VDM05] J. Vieira, F. M. Dias, and A. Mota. Hybrid neuro-fuzzy network-priori knowledge model in temperature control of a gas water heater system. *Proceedings of the 5th International Conference Hybrid Intelligent Systems (HIS'2005), São Paulo- Brasil, 06-09, November, 2005*.
- [Vie97] J. Vieira. *Desenvolvimento de Algoritmos de Controlo Baseados Em Lógica Difusa*. Dissertação de Mestrado, Departamento de Electrónica e Telecomunicações, Universidade de Aveiro, Portugal, 1997.
- [VM02] J. Vieira and A. Mota. Modeling and control of a water gas heater with neuro-fuzzy techniques. *Proceedings 3rd WSEAS FSFS International Conference on Fuzzy Sets and Fuzzy Systems, Interlaken, Switzerland*, 1:56–61, 2002.
- [VM03a] J. Vieira and A. Mota. Comparison of three neuro-fuzzy architectures in modelling a water gas heater: Input clustering data and input variables selection. *WSEAS Transactions on Systems*, 2(4):1005–1010, 2003.
- [VM03b] J. Vieira and A. Mota. Smith predictor based neural-fuzzy controller applied in a water gas heater that presents a large time-delay and load disturbances. *Proceedings IEEE International Conference on Control Applications, Istanbul, Turkey, 23 a 25 June*, 1:362–367, 2003.
- [VM04a] J. Vieira and A. Mota. Adaptive Smith predictive control of non-linear systems using neuro-fuzzy Hammerstein models. *Proceedings 1st International Conference on Informatics in Control Automation and Robotics, Setubal, Portugal*, 3:62–69, 2004.
- [VM04b] J. Vieira and A. Mota. Parameter estimation of non-linear systems with Hammerstein models using neuro-fuzzy and polynomial approximation approaches. *Proceedings of IEEE-FUZZ International Conference on Fuzzy Systems*, 2:849–854, 2004.



- [VM04c] J. Vieira and A. Mota. Parameter estimation of non-linear systems with Hammerstein models using neuro-fuzzy and polynomial approximation approaches. *Proceedings of IEEE-FUZZ International Conference on Fuzzy Systems, Budapest, Hungary*, 2:849–854, 2004.
- [VM04d] J. Vieira and A. Mota. Sistemas neuro-difusos: Técnicas de modelização não linear. *Revista do Departamento de Electrónica e de Telecomunicações da Universidade de Aveiro*, 4(2):251–258, 2004.
- [VM04e] J. Vieira and A. Mota. Water gas heater non-linear physical model: Optimization with genetic algorithms. *Proceedings of IASTED International Conference on Modelling, Identification, and Control*, 1:122–127, 2004.
- [VM04f] J. Vieira and A. Mota. Water gas heater non-linear physical model: Optimization with genetic algorithms. *Proceedings of IASTED International Conference on Modelling, Identification, and Control, Grindelwald, Switzerland*, 1:122–127, 23 a 25 February 2004.
- [VM05] J. Vieira and A. Mota. Combining first principles with grey-box approaches for modelling a water gas heater system. *Proceedings 13 th International Mediterranean Conference on Control and Automation, Limasol, Chipre*, 1:1131–1136, 2005.
- [Wan97] L. X. Wang. *A Course in Fuzzy Systems Ans Control*. Prentice Hall, New York, USA, 1997.
- [WH60] B. Widrow and Hoff. Adaptive switching circuits. *In 1960 IRE WESCON Convention Record - Reprinted in Neurocomputing - Foundations of Research eds. Anderson and Rosenfeld*, pages 96–104, 1960.
- [Wie58] N. Wiener. *Nonlinear Problems in Random Theory*. J. Wiley e Sons, The Tcchnology Press, M.I.T. New York, 1958.
- [Wig93] T. Wigren. Recursive prediction error identification using the nonlinear Wiener model. *Automatica*, 29:1011–1025, 1993.
- [WM92] L. Wang and J. M. Mendel. Fuzzy basis functions, universal approximation, and orthogonal least squares. *IEEE Transactions on Neural Networks*, 3(5):807–814, 1992.
- [XJ02] Qiang Xiong and A. Jutan. Grey-box modelling and control of chemical processes. *Chemical Engineering Science, Elsevier*, 57:1027–1039, 2002.
- [Zad65] L. A. Zadeh. Fuzzy sets. *Information and control*, 8:338–358, 1965.
- [Zad73] L. A. Zadeh. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on System Man and Cybernetics*, 3(1):28–44, 1973.

- [ZS94] X. Zhu and D. E. Seborg. Nonlinear predictive control based on Hammersstein models. *Proceedings Conference PSE'94, Kyongju, Korea*, 1994.
- [ZSK96] P. Zbicisk, P. Strumio, and W. Kamiski. Hybrid neural model of thermal drying in a fluidized bed. *Computers and Chemical Engineering*, 20, Supplement 1:S695–S700, 1996.

# Apêndice E

## Errata

Na página 62 no final do primeiro parágrafo falta um ponto final;

Na página 66 última linha, onde está “...  $y(k)$  é o vector das entradas,  $u(k)$  o é vector das saídas ...” deveria estar “...  $u(k)$  é o vector das entradas,  $y(k)$  o é vector das saídas ...” ;

Na página 69 no diagrama de blocos da figura 4.2, onde está  $g(.)$  deveria estar  $G(.)$ ;

Na página 79 segundo parágrafo terceira linha, onde está “... Os modelos híbridos baseiam-se ...” deveria estar “... Os modelos híbridos baseiam-se ...” ;

Na página 111 na legenda da figura 7.7, onde está “ ... (modelo simplificado) ...” deveria estar “... (sistema simplificado) ...”;

Na página 154 primeiro parágrafo quinta linha, onde está “... trás ...” deveria estar “... traz ...”;

Na página 157 desde a equação 8.43 até à página 158 equação 8.49, o texto que está deveria ser substituído por:

“... ”

$$MC_E \frac{d\Delta t(t)}{dt} = p_{aa}(t - \tau_{m1}(t)) - c_a(t - \tau_{m2}(t))C_E \Delta t(t) \quad (8.43)$$

em que  $M$  é a massa de água que se encontra dentro do permutador.

Escrevendo a equação 8.43 no domínio Laplaciano considerando  $\tau_{m1}(t) = \tau_{m1}$ ,  $\tau_{m2}(t) = \tau_{m2}$  e  $c_a(t - \tau_{m2}(t)) = C_A$  valores fixos e constantes obtêm-se a equação 8.44:

$$P_{aa}(s)e^{-\tau_{m1}s} = MC_E \Delta T(s)s + C_A C_E \Delta T(s) \quad (8.44)$$

$$\frac{\Delta T(s)}{P_{aa}(s)} = \frac{1}{MC_E s + C_A C_E} e^{-\tau_{m1}s} \quad (8.45)$$

em que  $P_{aa}(s)e^{-\tau_{m1}s}$  e  $\Delta T(s)$  são as transformadas de Laplace de  $p_{aa}(t - \tau_{m1})$  e  $\Delta t(t)$ , respectivamente.

Após alguma manipulação algébrica, a função de transferência do esquentador (relação entre a variação do aumento da temperatura à saída e potência térmica a absorver pela água) é apresentada na equação 8.46:

$$\frac{\Delta T(s)}{P_{aa}(s)} = \frac{\frac{1}{C_A C_E}}{\frac{M}{C_A}s + 1} e^{-\tau_{m1}s} \quad (8.46)$$

Rescrevendo a equação 8.46, como apresentado na equação 8.47, é possível efectuar a transformação desta função para o domínio discreto:

$$\frac{\Delta T(s)}{P_{aa}(s)} = \frac{\frac{1}{C_A C_E} \frac{C_A}{M}}{s + \frac{C_A}{M}} e^{-\tau_{m1}s} \quad (8.47)$$

Aplicando o operador  $e^{-sh} = z^{-1}$  à equação 8.47 e substituindo  $\tau_{m1}$  pela respectiva aproximação discreta  $d_1$ , obtém-se a equação 8.48:

$$\frac{\Delta T(z)}{P_{aa}(z)} = \frac{1}{C_A C_E} \frac{1 - e^{-\frac{C_A}{M}h}}{z - e^{-\frac{C_A}{M}h}} z^{-d_1} \quad (8.48)$$

Baseado na equação 8.48 e com  $h = 1$  s, e sabendo que  $d_1$ ,  $d_2$  e  $C_A$  variam no tempo, ou seja  $d_1(k)$ ,  $d_2(k)$  e  $c_a(k - d_2(k))$ , a equação 8.49 representa a função transferência do esquentador escrita do domínio discreto, cuja saída estimada é designada por  $\widehat{\Delta t}(k+1)$ :

$$\begin{aligned} \widehat{\Delta t}(k+1) &= \left( e^{-\frac{c_a(k-d_2(k))}{M}} \right) \Delta t(k) + \\ &\left( \frac{1}{c_a(k-d_2(k))C_E} \left( 1 - e^{-\frac{c_a(k-d_2(k))}{M}} \right) \right) p_{aa}(k - d_1(k)) \end{aligned} \quad (8.49)$$

...”;

Na página 174 quinto parágrafo, onde está “...  $e_{rf}(k) = \alpha e_r(k) + (1 - \alpha)(e_{rf}(k - 1))$  ...” deveria estar “...  $e_{rf}(k) = \alpha e_r(k) + (1 - \alpha)e_{rf}(k - 1)$  ...”;

Na página 191 primeiro parágrafo, onde está “... passou pela exploração vários métodos ...” deveria estar “... passou pela exploração de vários métodos ...”;

Na página 192 sétimo parágrafo, onde está “... área muito basta onde muito trabalho ...” deveria estar “... área muito vasta onde muito trabalho ...”;

# Apêndice A

## Lista de Publicações e Comunicações

### A.1 Artigos publicados em Jornal

#### A.1.1 Nacional

- J. Vieira e A. Mota - Sistemas Neuro-Difusos: Técnicas de Modelização Não Linear, *Revista Electrónica e Telecomunicações do Departamento de Electrónica e de Telecomunicações da Universidade de Aveiro*, 4(2):251-258, 2004.

#### A.1.2 Internacional

- J. Vieira e A. Mota - Parameter Estimation of Non-Linear Systems with Hammerstein Models Using Neuro-Fuzzy and Polynomial Approaches, *International Journal ELSEVIER, Engineering Applications of Artificial Intelligence* (submetido).
- J. Vieira, F. M. Dias e A. Mota - Neuro-Fuzzy Systems: A Survey, *WSEAS Transactions on Systems*, 2(3):414-419, 2004.
- J. Vieira, F. Dias e A. Mota - Artificial Neural Networks and Neuro-Fuzzy Systems in Model and Control: A Comparative Study, *International Journal ELSEVIER, Engineering Applications of Artificial Intelligence*, 3(17):265-273, 2004.
- J. Vieira e A. Mota - Comparison of Three Neuro-Fuzzy Architectures in Modelling a Water Gas Heater: Input Clustering Data and Input Variables Selection, *WSEAS Transactions on Systems*, 4(2):1005-1010, 2003.

## A.2 Artigos apresentados em Conferência

### A.2.1 Nacional

- J. Vieira, F. Dias e A. Mota - Control of a Kiln Using Neuro Fuzzy Techniques, *Proceedings 5th Portuguese Conference on Automatic Control, Aveiro, Portugal*, 05-07 Setembro, 1:519-524, 2002.

### A.2.2 Internacional

- J. Vieira, F. M. Dias e A. Mota - Hybrid Neuro-Fuzzy Network-Priori Knowledge Model in Temperature Control of a Gas Water Heater System, *Proceedings of the 5<sup>th</sup> International Conference on Hybrid Intelligent Systems (HIS'2005)*, Rio de Janeiro- Brasil, 06-09 November, 2005.
- F. M. Dias, A. Antunes, J. Vieira, A. Mota - Using the Levenberg-Marquardt for on-line training of a variant system, *International Conference on Artificial Neural Networks*, Spain, Setembro 2005.
- J. Vieira e A. Mota - Combining First Principles with Grey-Box Approaches for Modelling a Water Gas Heater System, *Proceedings of IEEE 13th Mediterranean Conference on Control and Automation*, Limassol, Cyprus, 27-29 June, 2005, 1:1131-1136.
- F. M. Dias, A. Antunes, J. Vieira, A. Mota - On-line training of Neural Networks: a sliding window approach for the Levenberg-Marquardt algorithm, *International Work-Conference on the Interplay Between Natural and Artificial Computation*, June 2005.
- F. M. Dias, A. Antunes, J. Vieira e A. Mota - Implementing the Levenberg-Marquardt Algorithm On-Line: A Sliding Window Approach with Early Stopping, *Proceedings of IFAC 2nd Workshop on Advanced Fuzzy/Neural Control*, Oulu, Finland, 16-17 September, 2004.
- J. Vieira e A. Mota - Adaptive Smith Predictive Control of Non-Linear Systems Using Neuro-Fuzzy Hammerstein Models, *ICINCO-IFAC Proceedings 1st International Conference on Informatics in Control Automation and Robotics*, Setúbal, Portugal, 25-28 August, 3:62-69, 2004.
- J. Vieira e A. Mota - Parameter Estimation of Non-Linear Systems With Hammerstein Models Using Neuro-Fuzzy and Polynomial Approximation Approaches, *Proceedings of IEEE-FUZZ International Conference on Fuzzy Systems*, Budapest, Hungary, 25-29 July, 2:849-854, 2004.
- J. Vieira, F. M. Dias e A. Mota - Neuro-Fuzzy Systems: A Survey, *Proceedings 5th WSEAS NNA International Conference on Neural Networks and Applications*, Udine, Italy, 25 a 27 March 2004.

- J. Vieira e A. Mota - Water Gas Heater Non-Linear Physical Model: Optimization with Genetic Algorithms, *Proceedings of IASTED International Conference on Modelling, Identification, and Control*, Gridelwald, Switzerland, 23 a 25 February, 1:122-127, 2004.
- J. Vieira e A. Mota - Neural-Fuzzy Model Versus Multi-Model Based Controllers Applied in a Water Gas Heater System, *Proceedings of IASTED International Conference on Artificial Intelligence and Applications*, Benalmádena, Spain, 08-10 September, 1:46-51, 2003.
- J. Vieira, F. Dias e A. Mota - Comparison Between Artificial Neural Networks and Neuro-Fuzzy Systems in Modelling and Control: A Case Study, *Proceedings 5th IFAC International Symposium on Intelligent Components and Instruments for Control Applications*, Aveiro, Portugal, 09-11 July, 1:271-27, 2003.
- J. Vieira e A. Mota - Smith Predictor Based Neural-Fuzzy Controller Applied in a Water Gas Heater that Presents a Large Time-Delay and Load Disturbances, *Proceedings IEEE International Conference on Control Applications*, Istanbul, Turkey, 23-25 June, 1:362-367, 2003.
- J. Vieira e A. Mota - Comparison of Three Neuro-Fuzzy Architectures in Modelling a Water Gas Heater: Input Clustering Data and Input Variables Selection, *Proceedings 4th WSEAS FSFS International Conference on Fuzzy Sets and Fuzzy Systems*, Athens, Greece, 29-31 May, 1:1005-1010, 2003.
- J. Vieira e A. Mota - Comparison of Two Neuro-Fuzzy Techniques in Modelling and Controlling a Kiln, *Proceedings IFAC International Conference on Intelligent Control Systems and Signal Processing*, Faro, Portugal, 08-11 April, 1:221-226, 2003
- J. Vieira e A. Mota - Modelling and Control of a water Gas Heater with Neuro-Fuzzy Techniques, *Proceedings 3rd WSEAS FSFS International Conference on Fuzzy Sets and Fuzzy Systems*, Interlaken, Switzerland, 11-15 February, 1:56-61, 2002.





# Apêndice B

## Glossário de Termos e Abreviaturas

### B.1 Abreviaturas

MIMO Multiple Input Multiple Output  
MISO Multiple Input Single Output  
SISO Single Input Single Output

PWM Pulse Width Modulation

RLMS Recursive Least Mean Square  
LMS Least Mean Square  
ARMA Auto Regressive Moving Average  
ARMAX Auto Regressive Moving Average with eXogenous inputs  
NARMAX Nonlinear Auto Regressive Moving Average with eXogenous  
NARX Nonlinear Auto Regressive with eXogenous inputs  
ARX Auto Regressive with eXogenous inputs  
OE Output Error  
BJ Box Jenkins

SVD Singular Value Decomposition  
QP Quadratic Programing  
ML Maximum Likelihood  
B&B Branch-and-bound  
ANFIS Adaptive Network Fuzzy Inference System  
ANN Artificial Neural Network  
AIC Akaike's Information Criterion  
GA Genetic Algorithms  
FARX Fuzzy Auto Regressive with eXogenous inputs

SA	Simulated Annealing
RLS	Recursive Least Square
PP	Pole Placement
ANN	Artificial Neural Networks
FS	Fuzzy Systems
ES	Expert Systems
RBF	Radial Based Functions
CS	Correlation Systems
FMNN	Feedforward Multilayer Neural Network
CART	Classification and Regression Tree

PID	Proportional Integral Derivative
PI	Proportional Integral
PD	Proportional Derivative
RAM	Random Access Memory
ROM	Read Only Memory

GPC	Generalized Predictive Control
DIC	Direct Inverse Control
AFC	Additive Feedforward Control
IMC	Internal Model Control
MPC	Model Predictive Control
GPC	Generalised Predictive Control
SPC	Smith Predictive Control
NMPC	Non-linear Model Predictive Control
LMPC	Linear Model Predictive Control
FHMP	Fuzzy Hammerstein Model-based Predictive Control

## B.2 Símbolos

### B.2.1 Genéricos

$t$	tempo
$k$	amostra índice $k$
$h$	período de amostragem
$\hat{x}(k)$	valor estimado da variável $x$ na iteração $k$
$r(k)$	signal de referência na iteração $k$
$e(k)$	erro entre duas variáveis na iteração $k$
$u$	sinais de entrada num sistema genérico
$y$	sinais de saída num sistema genérico

### B.2.2 Esquentador

$t_{aq}(t)$	temperatura da água quente no instante $t$
$t_{af}(t)$	temperatura da água fria no instante $t$
$\Delta t(t)$	aumento da temperatura da água à saída no instante $t$
$c_a(t)$	caudal da água quente no instante $t$
$f_g(t)$	fluxo de gás no instante $t$
$\tau_{m1}(t)$	tempo morto total relativo ao fluxo de gás no instante $t$
$\tau_{m0}(t)$	tempo morto constante relativo ao fluxo de gás no instante $t$
$\tau_{m2}(t)$	tempo morto relativo ao caudal de água no instante $t$
$\tau_s(t)$	tempo de subida no instante $t$
$\tau_d(t)$	tempo de descida no instante $t$
$p_e(t)$	perdas para o exterior no instante $t$
$e_s(t)$	energia do sistema no instante $t$
$p_{ffg}(t)$	potência térmica fornecida pelo fluxo de gás no instante $t$
$p_p(t)$	potência térmica perdida para o exterior no instante $t$
$p_{af}(t)$	potência térmica que entra no sistema com a água fria no instante $t$
$p_{aq}(t)$	potência térmica que sai do sistema através da água quente no instante $t$
$p_{aa}(t)$	potência térmica absorvida pela água no instante $t$
$C_E$	calor específico da água
$M$	massa de água dentro da câmara de permutação
$PE$	potência estática
$PE_{MAX}$	potência estática máxima
$PE_{MIN}$	potência estática mínima
$d_1(k)$	tempo morto discreto total relativo ao fluxo de gás na iteração $k$
$d_2(k)$	tempo morto discreto relativo ao caudal de água na iteração $k$
$t_{aq}(k)$	temperatura da água quente na iteração $k$
$t_{af}(k)$	temperatura da água fria na iteração $k$
$\Delta t(k)$	aumento da temperatura da água à saída na iteração $k$
$c_a(k)$	caudal da água quente na iteração $k$
$f_g(k)$	fluxo de gás na iteração $k$
$r(k)$	sinal de referência para a temperatura da água desejada na iteração $k$

$\mathbf{D}^{ND}$	vector de dados de treino dinâmicos, $ND$ amostras
$\mathbf{H}^{ND}$	vector de dados de teste dinâmicos, $NH$ amostras
$\mathbf{E}^{NE}$	vector de treino estáticos e dinâmicos (caudal fixo), $NE$ amostras
$\mathbf{G}^{NG}$	vector de treino estáticos e dinâmicos (fluxo de gás fixo), $NG$ amostras
$\mathbf{EE}^{NEE}$	vector de dados de treino estáticos (caudal fixo), $NEE$ amostras
$\theta$	vector de parâmetros a determinar

### B.2.3 Modelos Neuro-Difusos

$R_i$	regra $i$
$A_{ij}$	conjunto difuso relativo às entradas $i$ e $j$
$\mu_A$	grau de pertença relativamente ao conjunto difuso $A$
$\beta_i$	grau de pertença global relativo à regra $i$
$\bar{\beta}_i$	grau de pertença global normalizado relativo à regra $i$

### B.2.4 Modelos Hammerstein

$f(\cdot)$	função não linear estática
$G(\cdot)$	modelo linear dinâmico
$u(k)$	signal de entrada na iteração $k$
$z(k)$	signal intermédio entre $f(\cdot)$ e $G(\cdot)$ na iteração $k$
$y(k)$	signal de saída na iteração $k$
$\bar{y}(k)$	signal de saída em estado estacionário na iteração $k$
$\bar{u}(k)$	signal de entrada em estado estacionário na iteração $k$

### B.2.5 Modelos Híbridos

$x_i$	variável determinada através das equações de balanço do sistema
$f(\cdot) g(\cdot)$	funções que representam as equações diferenciais do sistema
$f_{CCN}(\cdot)$	função do tipo caixa cinzenta ou negra

## Apêndice C

# Sistema de Controlo do Esquentador - *Hardware*

Este anexo apresenta o esquema eléctrico completo da placa de controlo utilizada na aquisição de dados e na realização dos testes de avaliação dos vários controladores propostos.

O esquema está subdividido em quatro páginas:

- Na primeira página encontram-se os esquemas eléctricos relativos ao microcontrolador 89C51, aos dois dígitos de sete segmentos, aos botões de pressão para o interface com o utilizador e são ainda apresentados aos circuitos de condicionamento de sinal para os sensores de temperatura.
- Na seguinte página são apresentados os esquemas eléctricos relativos ao microcontrolador PIC16C84, à electrónica de interface com a válvula de gás *on/off* e aos circuitos de condicionamento de sinal para a medida das grandezas: caudal de água, pressão e excesso de temperatura.
- Na terceira página estão definidos os esquemas eléctricos de interface com o ventilador e com o faiscador e o circuito de condicionamento de sinal do circuito que permite verificar a ocorrência de ionização.
- Na última página encontram-se os esquemas eléctricos de interface com a válvula de gás proporcional e os circuitos de interface para o estabelecimento da comunicações via RS232 com o computador pessoal.



## Apêndice D

# Sistema de Controlo do Esquentador - *Software*

O microcontrolador utilizado para o desenvolvimento do protótipo do esquentador controlado automaticamente foi o 89C51RD da PHILIPS. A linguagem de programação utilizada foi o ANSI C utilizando a ferramenta de desenvolvimento IAR Embedded Workbench® versão 3.1.

Embora este trabalho não tivesse como prioridade a execução dos algoritmos de controlo desenvolvidos no próprio microcontrolador, estes foram aí testados observando os tempos de execução e os espaços em memória RAM e ROM necessários para cada algoritmo.

A tabela D.1 apresenta os tempos de execução e os espaços de memória RAM e ROM necessários para as quatro malhas de controlo quando executadas pelo microcontrolador.

Código Programa	Memória ROM	Memória RAM	Tempos de Execução
	[bytes]	[bytes]	[mili-segundos]
Sem Controlador	7970	144	–
CSPL	8052	152	15
CSPHS	8944	184	32
CSPHA	9362	220	65
CSPND	12542	321	324

Tabela D.1: Tempos de execução e espaço em memória necessários para os quatro algoritmos de controlo.

Este estudo confirma que os tempos de execução aumentam com o aumento da complexidade dos algoritmos de controlo/modelo utilizados. Executados num microcontrolador as diferenças de complexidade destes algoritmos são ainda mais visíveis devido ao seu baixo poder de cálculo.

Como se pode ver pela tabela D.1 existe uma parte do código que resulta do código necessário para as restantes funções do programa global.

No CSPL são executados simplesmente algumas multiplicações que são executadas num tempo relativamente baixo sendo o espaço de memória ROM e RAM necessário igualmente baixo. No CSPND são executadas muitas funções exponenciais em formato virgula flutuante que implicam um aumento exponencial do tempo de execução e do espaço de memória RAM e ROM necessários.

Pode-se concluir que todos os algoritmos de controlo testados podem ser executados pelo microcontrolador uma vez que, o período de amostragem de 1 segundo é suficiente para a execução de todo o código de controlo para qualquer algoritmo.